Interior structure of the Moon – constraints from seismic tomography, gravity and topography

Bernhard Steinberger\textsuperscript{a, b, *}, Dapeng Zhao\textsuperscript{c}, Stephanie C. Werner\textsuperscript{b}

\textsuperscript{a}GFZ German Research Centre for Geosciences, Telegrafenberg, 14473 Potsdam, Germany
\textsuperscript{b}Centre for Earth Evolution and Dynamics, University of Oslo, PO Box 1028, 0315 Oslo, Norway
\textsuperscript{c}Department of Geophysics, Tohoku University, Sendai 980-8578, Japan

Abstract

Seismic tomography can be combined with constraints from geoid, topography and other surface observations to gain information about mantle structure and dynamics. This approach has been taken with much success for the Earth mantle, and here it is, for the first time, applied to the Moon. Lunar tomography has much lower resolution as for the Earth and is mostly restricted to the near side, nevertheless we can assess under what assumptions the fit between predicted geoid (based on a tomography model) and observed geoid is best: Among the models tested, we find the most similar pattern (correlation about 0.5) if we only consider tomography below 225 km depth, if density anomalies cause little or no dynamic topography and if we compare to the geoid with the flattening \((l = 2, m = 0)\) term removed. This could mean that (a) like for the Earth, seismic anomalies shallower than 225 km are caused by a combination of thermal and compositional effects and therefore cannot be simply converted to density anomalies; (b) the lithosphere is sufficiently thick to prevent dynamic topography more than a small fraction of total topography; and (c) flattening is a “fossil” bulge.
unrelated to present-day mantle anomalies. However, we have to be cautious with interpreting our results, because for models with comparatively higher correlation and a conversion from seismic velocity to density anomalies similar to the Earth’s upper mantle, the amplitude of the predicted geoid is much lower than observed: This could either mean that the tomography model is strongly damped, or that the geoid is mostly due to shallow causes such as crustal thickness variations, with only a small part coming from the deeper mantle.

**Keywords:** Moon interior structure, seismic tomography, gravity anomalies

1. **Introduction**

Seismic tomography provides a powerful tool to gain information about the interior of the Earth, in particular if it is interpreted jointly with gravity and topography. This was first attempted in the 1970s (Dziewonski et al., 1977), and by now, tomography of the Earth’s mantle has proliferated and led to countless publications. Also in the 1970s, seismometers installed during four of the Apollo missions (Fig. 1) recorded seismograms. Yet only recently this seismic information has been utilized to construct a lunar tomography model (Zhao et al., 2008, 2012). Even the existence of a lunar core has only recently been proven (Weber et al., 2011). We have thus reached a stage in learning about the lunar interior comparable to where we were regarding the Earth interior in the 1970s. Whereas for all other planets we still have at most gravity and topography information, the Moon now is the only other planetary body besides Earth, where we can jointly utilize information from seismic tomography, gravity and topography. This paper represents a first
attempt to do so.

Also recently, improved models of lunar gravity (Araki et al., 2009; Konopliv et al., 2013) and topography (Namiki et al., 2009) have been released. Topography and gravity equipotential surface are shown in Fig. 1 A and B. Although the term “geoid” etymologically refers to the Earth, we will use it here also for the gravity equipotential surface of the Moon to follow common practice, although, in analogy “selenoid” would be more appropriate.

A feature that has been noted early on and that is clearly evident in Fig. 1 (B) are geoid highs associated with five nearside ringed maria (Imbrium, Serenitatis, Crisium, Nectaris, and Humorum). These have been attributed to mass concentrations or mascons (Muller and Sjogren, 1968) that exist beneath the center of all of them. Here we would like to investigate possible sources of gravity anomalies in the deep interior of the Moon, and therefore attempt to remove the effect of mascons. This is done in Fig. 1 C and D where we have interpolated geoid and topography inside the mascons from surrounding values.

Another notable feature is the flattening of the lunar geoid which is, to its largest part, non-equilibrium, as the Moon is now rotating very slowly. It has been suggested to represent a fossil shape frozen into the lithosphere early in its orbital evolution (Jeffreys, 1976; Lambeck and Pullan, 1980). However, it may also be merely a consequence of internal density anomalies, and the fact that any planetary body always orients itself relative to its spin axis such that geoid highs are close to its equator (the minimum energy configuration for a synchronously rotating satellite, e.g. Lambeck (1988)), although these density anomalies and shape may also be a “fossil” remains from a previous
Figure 1: Caption on separate page.
Figure 1: (A): Lunar topography (Namiki et al., 2009) relative to the geoid. Triangles indicate Apollo seismometer locations. Procellarum KREEP terrane (Wieczorek and Phillips, 2000) is outlined in black. Following Laneuville et al. (2013), we use the 4 ppm Thorium abundance contour to define the KREEP outline. High-altitude abundances are adopted from Lawrence et al. (2000), online at http://www.lunar.lanl.gov/pages/GRSthorium.html. Also shown are the distribution of mare units (white, after Werner and Medvedev, 2010) that fill the large impact basins with basaltic material mostly on the near-side of the Moon. Map projection centered on near side. (B): Lunar geoid (Araki et al., 2009) relative to a sphere. Other features as in (A). (C): Near-side topography with depressions associated with mascons removed. The five mascons considered are shown as circles. At grid points inside circles, topography is initially set to zero, and iteratively replaced by the mean of values at the four neighbouring grid points until, after 1000 iterations, convergence has been approximately achieved. In this way, topography above mascons is interpolated from surrounding values. (D): Near-side geoid after the effect of mascons has been removed in an analogous manner to (C).
convection state (Matsuyama, 2012).

Regions of low topography generally coincide with mare units (Fig. 1). The relatively younger mare units in the west (Hiesinger et al., 2011) lie within a region known as Procellarum KREEP terrane (Wieczorek and Phillips, 2000; Grimm, 2013) which has been suggested to be underlain by hotter than average mantle that could also be responsible for the relatively recent volcanism until $\approx 1 \text{ Gyr}$ (Hiesinger et al., 2011) or even younger (Braden et al., 2014).

The relation of internal density anomalies and geoid depends on whether the lunar mantle is still convecting, and if so, at what depths. Although the Moon is geologically “dead” with its surface preserved for billions of years, therefore presumably has a thick rigid outer shell, it is possible that convection is still ongoing in its deep interior (Turcotte and Oxburgh, 1970; Meissner, 1977; Schubert et al., 1977).

In this paper, we first present spectral characteristics of lunar geoid and topography. Our analysis in section 2 is mostly not new, but mainly meant to show that there are indications for both a deep and a shallow origin of lunar geoid undulations. Hence in this way we motivate and set the stage for the new work (at least new for the Moon) combining information from seismic tomography, geoid and topography to learn more about the interior of the Moon. Seismic tomography, the moonquake data it is based on, and possible inferences on lunar internal density structure are discussed in section 3, and how such density anomalies relate to geoid and topography in section 4. Because there are rather large uncertainties in (i) the seismic velocity anomalies (ii) conversion to density anomalies (iii) elastic lithosphere thick-
ness, hence how internal density anomalies relate to topography and geoid, we will make certain approximations which we think are justified based on the low level of accuracy we can expect to achieve. Because many of the uncertainties are also hard to quantify, we will not attempt a formal error analysis. Rather, we will use the approach that – for the same reasons – is common in geodynamic modelling of the Earth mantle: That we vary certain parameters and assumptions within a range that appears reasonable based on what we know, and compare results with observations available. In this way, we expect to find out which parameters and assumptions are most suitable to explain available data.

2. Lunar geoid and topography: spectral characteristics and correlations

Gravity and topography, as well as density anomalies, can be expressed in terms of spherical harmonic coefficients, e.g. the gravity potential \( U \) on a spherical surface with the lunar radius \( r_0 \) can be expressed as

\[
U = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} U_{lm} Y_{lm}
\]  

(1)

whereby \( Y_{lm} \) are “fully normalized” spherical harmonic functions – see e.g. Wieczorek (2007). Gravity potential is usually expressed in terms of dimensionless coefficients \( C_{lm}^G \) and \( S_{lm}^G \), i.e. coefficients \( U_{lm} \) are normalized by dividing through \(-GM/r_0\), whereby \( G \) is the Newtonian constant of gravitation and \( M \) is the total mass: \( U_{lm} = (-GM/r_0) \cdot C_{lm}^G \) for \( m \geq 0 \) and \( U_{lm} = (-GM/r_0) \cdot S_{|lm|}^G \) for \( m < 0 \), \( C_{00}^G = 1 \), \( C_{10}^G = C_{11}^G = S_{11}^G = 0 \).
 Whereas in case of the Earth the $C_{20}^G$ coefficient is largely due to equilib-
rium flattening and hence the geoid is defined relative to a reference ellipsoid,
this is not the case for the Moon, because it rotates much more slowly. Using
the Darwin-Radau equation, one can verify that the equilibrium value of $C_{20}^G$
is only a small fraction (of the order of 1%) of the observed coefficient, hence
we do not correct for it and use a sphere for reference shape.

 The power spectrum of geoid and topography, i.e. power as a function of
spherical harmonic degree, provides further information about the interior.
Based on Hipkin (2001) we define average geoid power of spherical harmonic
degree $l$ in terms of these dimensionless coefficients as

$$\langle P_l^G \rangle = r_0^2 \cdot (l + 1) \left( C_{l0}^G \cdot \sum_{m=1}^{l} (C_{lm}^G \cdot S_{lm}^G) \right)$$  \hspace{1cm} (2)

The definition of average topography power $\langle P_l^T \rangle$ is entirely analogous. Fig.
2 shows $\sqrt{\langle P_l^T \rangle}$ and $\sqrt{\langle P_l^G \rangle}$, (in units of meters), the geoid-topography ra-
tio $\sqrt{\langle P_l^G \rangle} / \langle P_l^T \rangle$ and the geoid-topography correlation for each spherical
harmonic degree $l$.

 The square root of geoid power (blue line in Fig. 2 A) generally decreases
with increasing degree. Above degree 15 it approaches the dotted line $\sim$
$[r_0/(r_0-30 \text{ km})]^l$, which becomes a “white” spectrum (constant power) when
it is downward-continued to depth 30 km. After the effect of mascons is
removed, power is somewhat reduced, particularly in the degree range $9 \leq$
$l \leq 14$. The square root of the remaining power (red line) in the degree
range 3 to 10 approximately follows the dashed line $\sim [r_0/(r_0-300 \text{ km})]^l$,
which becomes a “white” spectrum when downward-continued to depth 300
km. For degree 10 and above, it approximately follows the dotted line. The
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Parameter</th>
<th>Value</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>$G$</td>
<td>constant of gravitation</td>
<td>$6.674 \cdot 10^{-11} \text{m}^3/\text{kg}/\text{s}^2$</td>
<td>Taylor and Mohr (2011)</td>
</tr>
<tr>
<td>$M$</td>
<td>mass</td>
<td>$7.3463 \cdot 10^{22} \text{kg}$</td>
<td>$GM$ from Konopliv et al. (2013)</td>
</tr>
<tr>
<td>$r_0$</td>
<td>radius</td>
<td>1737.1 km</td>
<td>Smith et al. (1997)</td>
</tr>
<tr>
<td>$\bar{\rho}$</td>
<td>average density</td>
<td>3344 kg/m$^3$</td>
<td>$M/(r_0^3 \cdot 4\pi/3)$</td>
</tr>
<tr>
<td>$\rho_c$</td>
<td>crustal density</td>
<td>2900 kg/m$^3$</td>
<td>Wieczorek et al. (2006), Tbl. 3.10</td>
</tr>
<tr>
<td>$t_c$</td>
<td>crust thickness</td>
<td>50 km</td>
<td>Wieczorek et al. (2006), Tbl. 3.10</td>
</tr>
<tr>
<td>$t_e$</td>
<td>elastic lithosphere thickness</td>
<td>65, 122 or 240 km</td>
<td></td>
</tr>
<tr>
<td>$t_l$</td>
<td>thermal lithosphere thickness</td>
<td>240 km</td>
<td></td>
</tr>
<tr>
<td>$\rho_m$</td>
<td>uppermost mantle density</td>
<td>3310 kg/m$^3$</td>
<td>see Appendix</td>
</tr>
<tr>
<td>$r_b$</td>
<td>core radius</td>
<td>330 km</td>
<td>Weber et al. (2011)</td>
</tr>
<tr>
<td>$E$</td>
<td>Young’s modulus</td>
<td>$6.5 \cdot 10^{10} \text{Pa}$ ($1.6 \cdot 10^{11} \text{Pa}$)</td>
<td>Turcotte et al. (1981)</td>
</tr>
<tr>
<td>$\nu$</td>
<td>Poisson’s ratio</td>
<td>0.25</td>
<td>Turcotte et al. (1981)</td>
</tr>
<tr>
<td>$T_0$</td>
<td>surface temperature</td>
<td>253 K</td>
<td>Williams (2010)</td>
</tr>
<tr>
<td>$T_b$</td>
<td>CMB temperature</td>
<td>1687 K (adiabatic)</td>
<td></td>
</tr>
<tr>
<td>MOI</td>
<td>moment of inertia factor</td>
<td>0.3932</td>
<td>Konopliv et al. (1998)</td>
</tr>
</tbody>
</table>
Figure 2: Caption on separate page
Figure 2: (A): Square-root of lunar topography power $\sqrt{\langle P^T_l \rangle}$ and geoid power $\sqrt{\langle P^G_l \rangle}$ (Eq. 2). Black and blue lines correspond to Fig. 1 A and B, red corresponds to Fig. 1 C and D. Between $l=2$ and 3, orange lines are obtained by replacing $C^G_{20}$ with zero, i.e. not considering excess flattening, violet lines by reducing it to one-third of its value. The dark green line is for residual gravity corresponding to the violet (and red, for $l \geq 3$) line, if the effect of topography, using crustal density 2900 kg/m$^3$ and assuming isostatic compensation at depth 70 km, is subtracted. The black dotted line is a white spectrum upward continued from 30 km depth, $P_{30\text{km}} = 42 \cdot [r_0/(r_0 - 30 \text{ km})]^l$. The shaded area shows $P_{30\text{km}} \cdot (1 \pm 1/\sqrt{2l + 1})$, a crude estimate for the expected standard deviation, if coefficients were picked at random from a normal distribution (Steinberger et al., 2010). The black dashed line is a white spectrum upward continued from 300 km depth, $P_{300\text{km}} = 173 \cdot [r_0/(r_0 - 300 \text{ km})]^l$, with corresponding expected standard deviation as dark shaded area. (B): Geoid-topography ratio $\sqrt{\langle P^G_l \rangle / \langle P^T_l \rangle}$. The bright green line is the theoretical geoid-topography ratio for the correction applied for the green line in (A). (C): Geoid-topography correlation. The thick pink line is for only the near side, whereby for each spherical harmonic degree $l$ a spectral window from $l-5$ to $l+5$ is considered. The brown line is the corresponding curve for the far side, the dashed line for the whole Moon.
square root of degree-two power lies considerably above the dashed line, but this is largely due to the large “flattening” coefficient $C_{20}^G$ which may at least partly represent a “frozen” equilibrium shape from an early period: If $C_{20}^G$ is set to zero – which corresponds to assuming it entirely represents a fossil bulge – the square root of the remaining degree-two power is actually below the dashed line and below the degree 3 value. If it is set to one third of its actual value (approximately equal to $C_{22}^G$; other degree-two coefficients are much smaller) – which corresponds to assuming that the other two thirds represent a fossil bulge – the resulting spectrum (violet line) closely follows the dashed line in the whole range $l = 2$ to 10.

Geoid-topography correlation (blue line in Fig. 2 C) is generally positive, but less so (mostly below 0.5) in the degree range $\approx 10$-30, and even negative for degrees 10 and 11. We expect that this reduced correlation can be attributed to the mascons, which are associated with positive geoid and negative topography, hence contribute a negative geoid-topography correlation. Accordingly, removing the effect of mascons increases correlation, in particular in this mid-degree range $\approx 10$-30. Correlation for degrees 10-30 remains somewhat lower (around 0.6) than it is for lower and higher degrees (around 0.8 for $l = 3$ to 9 and above $l = 30$). The remaining reduction in correlation in this degree range is mostly caused on the near side. If we separately consider correlation on the near and far side for a spectral window from $l - 5$ to $l + 5$, we find a drop to values between 0.2 and 0.3 for $l = 15$ to 22 on the near side (thick pink line), whereas on the far side, correlation remains above 0.65 (brown line).

Degree-two correlation is low (around 0.2) if the excess geoid flattening
is removed, but much higher (≈ 0.6) if the $C_{20}^G$ coefficient is kept or reduced to one third of its value. This occurs because, apart from the $l = 2, m = 0$ flattening term, degree-two geoid and topography are poorly correlated, whereas both geoid and topography (relative to the geoid) show an excess flattening (positive coefficient for $l = 2, m = 0$).

After the effect of mascons has been removed, geoid-topography ratio is rather constant around 0.05 to 0.08 for $l=3$ to 11. It is somewhat larger for degree 2, but drops to the same range if $C_{20}^G$ is again reduced to one third of its original value (i.e. assuming the remaining part is “frozen” excess flattening).

A geoid-topography ratio that only weakly depends on $l$ along with high geoid-topography correlation in the degree range $l = 3$ to 9 can be explained by isostatically compensated topography: If topography of a crust with density $\rho_c$ is compensated at depth $z_0$, the geoid-topography ratio is

$$GTR = \frac{1 - \left(1 - \frac{z_0}{\rho_c}ight)^l}{\frac{(2l+1)}{3\bar{\rho}} - 1}$$

where $\bar{\rho}$ is the average density of the Moon. This will be derived in the section 4. With $\rho_c/\bar{\rho} = 0.867$ (corresponding to $\rho_c=2900$ kg/m$^3$) and $z_0 = 70$ km the geoid-topography ratio (green line Fig. 2 B) becomes very similar to the observation-based ratio with the effect of mascons removed (red line).

The ratio in Eq. 3 is reduced, if a lower value for crustal density (Wieczorek et al., 2013) is used. If we subtract the gravity due to isostatically compensated crust we obtain much lower “residual” power (compare green line to red line in Fig. 2 A), particular in the degree range where geoid-topography correlation is high (up to degree 9 and above degree 30). This residual power
remains very similar if the lower value for crustal density from Wieczorek et al. (2013) is used. Here a compensation depth $z_0 = 70 \text{ km}$ was chosen, because residual power for low degrees (up to $l=9$) reaches its minimum, as a function of $z_0$, at approximately that value. For the high degrees (above $l=30$) results are essentially indistinguishable from uncompensated topography. For even higher degrees (above 80, and outside the range we discuss here), Zuber et al. (2013) show that the gravity signal is to almost 100% associated with the topography, or near surface crustal density variations as to be expected. For the low degrees on which we focus here, the gravity field is equally well represented by Araki et al. (2009) and Konopliv et al. (2013), and it is therefore not necessary to update to the later model.

On the other hand, we have seen that the geoid power spectrum after removing the effect of mascons can be made approximately “white” for degrees $l \geq 10$ through downward continuation to depth 30 km. In contrast, for degrees 2 to 10, with appropriate adjustment of the $C_{20}^{G}$ term, it becomes approximately “white” through downward-continuation to depth 300 km. The depth of downward continuation in order to make the power spectrum white can be seen as an indication of the source depth of the gravity anomalies (Hipkin, 2001; Steinberger and Holme, 2002; Steinberger et al., 2010). Hence we can surmise that gravity anomalies for $l \geq 10$ mostly originate at a shallow depth $\approx 30 \text{ km}$, and only those at $l < 10$ may originate at a larger depth $\approx 300 \text{ km}$. Due to the limited resolution of the tomography model, this work focusses on low degrees for which a mass anomaly depth around 300 km is suggested. We note that the red line in Fig. 2 is approximately fit by the dashed line up to and including $l = 10$. But since the red line at $l = 10$ is
below the dotted line, we treat in the following only degrees up to $l = 9$ as possibly related to deeper density anomalies.

There is no physical law that demands a white spectrum at source depth, but interestingly, the high-degree end of all terrestrial planets and the Moon for which sufficiently high-resolution gravity data are available can be fit by a straight line that corresponds to a white spectrum when downward-continued to a rather shallow (lithospheric) depth (Fig. 3). So there is at least some reason to believe that the slope of the geoid spectrum indicates source depth, and we have an apparent paradox that the geoid-topography ratio and correlation indicates a shallow gravity source, whereas the geoid spectrum rather indicates a deeper source for low degrees.

In the degree range 2-9, where density anomalies causing geoid undulations are possibly located deeper than the crust, regional differences in geoid-topography correlation and ratio exist (Fig. 4): In the northern part of the near side, where most of the dark lunar maria occur, and which also contains the Procellarum KREEP terrane (Wieczorek and Phillips, 2000; Grimm, 2013), geoid-topography correlation is generally lower (about 0.4-0.6) than elsewhere (around 0.8). Also, in this region, geoid-topography ratio is higher (around 0.1) than elsewhere (around 0.06-0.08). Both indicates that in this region, isostatic compensation at crustal levels is less dominant as a cause of geoid undulations. This could mean that in this region, which is also similar to the region where seismic coverage is best, due to distribution of seismometers, the contribution of “deep” gravity sources is larger than elsewhere. Here we resort to seismic tomography in order to contribute to a resolution of the issue of deep vs. shallow gravity sources.
Figure 3: Geoid spectra of Earth, Venus, Mars and Moon in comparison. Also shown is a straight-line fit for the high-degree end, and the depth of downward-continuation that would make the line horizontal. The difference in degree above which this straight line approximately fits is probably related to the different size of these bodies: For smaller bodies, the same distance near the surface corresponds to lower degree. Also, smaller bodies have probably cooled more and hence have a thicker lithosphere, where this straight-line-fit part of the spectrum is thought to originate. See Steinberger et al. (2010) for data sources for Mars, Venus and Earth, and further analysis of their spectra.
Figure 4: Regional geoid-topography correlation (A) and ratio (B) for spherical harmonic degrees $l=2-9$, after the effect of mascons has been removed. Excess flattening has been reduced by changing $C_{20}^G$ to one-third of its value, corresponding to the violet line in Fig. 2. Correlations and ratios have been computed for the hemisphere (90-degree cap) centered on each grid point. Triangles indicate seismometer locations. Procellarum KREEP terrane is outlined in black, mare units are outlined in white.
3. Models of density anomalies in the lunar mantle

We infer internal density anomalies from the tomography model of Zhao et al. (2012), which is derived from seismograms recorded in the 1970s by seismometers installed during four of the Apollo missions. The worst Apollo arrival-time data could have picking errors up to tens of seconds, but a large fraction of the Apollo data including $\approx 7000$ deep moonquakes are still very good (Nakamura, 2005). A very best set of the Apollo data containing about 100 best-located moonquakes was selected very carefully to determine the lunar tomography model (Zhao et al., 2008, 2012). Because of the damping and smoothing regularizations applied to the tomographic inversion, the maximum velocity perturbation of the tomographic model is $\approx 1.5\%$ (see Fig. 5), whereas the uncertainty of the velocity perturbations is estimated to be less than 0.2\%.

Given the small number and limited distribution of seismometers, the model is of low resolution and not global, and anomalies are set to zero where there is no seismic ray coverage. We nevertheless expand the model in spherical harmonics globally. The model is given in layers at depths of 20 km, 150 km, 300 km, 500 km, 700 km, 900 km, 1100 km and 1300 km. We assume the layer boundaries at the midpoints between layer depths, and accordingly assign thickness 85 km, 140 km, 175 km, 200 km, 200 km, 200 km, 200 km, and 207.1 km to these layers. The lowermost layer extends to the core-mantle boundary (CMB) for which we use a depth 1407.1 km (Weber et al., 2011). The model in its spherical harmonic expansion is shown in Fig. 5. Comparison with the figure given in Zhao et al. (2012) shows that the spherical harmonic expansion represents the model well where it
is constrained by data, smoothly approaches zero elsewhere, and does not introduce artifacts.

We convert relative seismic velocity anomalies \( \delta v_s/v_s \) to density anomalies \( \delta \rho/\rho \) through \( \delta \rho/\rho = C \delta v_s/v_s \). Such a conversion with a constant or depth-dependent \( C \) is often used when interpreting tomography on Earth, e.g. assuming that both seismic velocity and density anomalies are caused by temperature anomalies. For the Earth, detailed mineral physics models allow in this case to compute \( C \) as a function of depth. For the upper mantle (<400 km) pressure range, which includes the pressure range of the lunar mantle, \( C \) remains nearly constant \( \approx 0.22 \). This is e.g. derived by Steinberger and Calderwood (2006), based on previous work. Assuming the mineralogy of the lunar mantle is similar to the Earth’s mantle, and given the large uncertainties of our model, using a constant conversion factor should hence be an appropriate approximation.

Since Zhao et al. (2012) note that most deep moonquakes occur in areas with average to higher velocity or at the boundary between high- and low-velocity zones, we tentatively also design a density model that is only based on “deep” moonquake locations (Nakamura, 2005) below 225 km depth (Fig. 5). Using the same depth layers, we assign a constant positive density anomaly to a block extending 22 degrees in both latitude and longitude in one layer around each quake location. The block size 22 degrees was chosen such that the volume with positive anomaly becomes rather continuous, because we expect that many gaps between moonquake locations are due to the short recording period, and moonquake locations would be more closely spaced over longer time periods. At a depth of 900 km, around which most
Figure 5: Spherical harmonic expansion up to degree 31 of the Zhao et al. (2012) tomography model. Triangles indicate Apollo seismometer locations. Circles indicate moonquake locations (Nakamura, 2005) within each depth layer. For each layer, the region where the tomography model is actually constrained by data is outlined in black.
moonquakes occur, 22 degrees corresponds to 321 km. If blocks around different moonquakes overlap, density anomalies are not added, i.e. the density model is “binary” in that only the value zero and a single constant positive value are possible.

4. Relation of gravity, topography and density anomalies

Here we provide a general outline of how gravity anomalies are related to density anomalies and the topography of interfaces (including the surface), and how for different modelling assumptions topography is in turn related to density anomalies. We will use the kernel formalism to describe this relation, and apply and simplify this approach for the Moon.

Internal density anomalies $\delta \rho$ at radius $r$ expanded in terms of spherical harmonic coefficients $\rho_{lm}(r)$ cause coefficients $U_{lm,1}$ of the gravity potential at the surface radius $r_0$

$$U_{lm,1} = -\frac{GM}{r_0^2} \cdot \frac{3}{(2l + 1)\bar{\rho}} \int_{r_b}^{r_0} \rho_{lm}(r) \cdot \left( \frac{r}{r_0} \right)^{l+2} \, dr$$

(4)

whereby $r_b$ is the core radius of the moon (see Table 1). Similarly, an interface at radius $r_i$ with density contrast $\Delta \rho$ and topography expansion coefficients $h_{lm}$ (relative to spherical shape) results in gravity potential expansion coefficients

$$U_{lm,2} = -\frac{GM}{r_0^2} \cdot \frac{3 \cdot \Delta \rho}{(2l + 1)\bar{\rho}} \cdot h_{lm} \cdot \left( \frac{r_i}{r_0} \right)^{l+2}$$

(5)

in the approximation that the topography is small relative to radius. Possible interfaces with density contrast include the core-mantle boundary, the boundary between crust and mantle and the surface. However, since $(r_b/r_0)^{l+2} = 0.0013$ for $l = 2$, and it further decreases with increasing $l$, we neglect gravity
anomalies due to core-mantle boundary topography. Concerning topography possibly caused by internal mantle density anomalies, we shall assume that topography at the crust-mantle interface is identical to surface topography (i.e. crustal thickness is not affected by topography due to mantle density anomalies). Given the low resolution of the tomography model, any density anomalies inferred from the tomography model and topography caused by these density anomalies will be long-wavelength (small \( l \)). Therefore we will replace the combined effect of topography at the surface (radius \( r_0 \)) with density contrast \( \rho_c \) and at the crust-mantle boundary (radius \( r_0 - t_c \)) with density contrast \( \rho_m - \rho_c \) by topography at the surface with density contrast \( \rho_m \). The relative error made through this approximation is

\[
\frac{\rho_m - \rho_c}{\rho_m} - \frac{\rho_m - \rho_c}{\rho_m} \cdot \left( \frac{r_0 - r_0}{r_0} \right)^{l+2} = \frac{\rho_m - \rho_c}{\rho_m} \left( 1 - \left( \frac{r_0 - t_c}{r_0} \right)^{l+2} \right) \approx \frac{\rho_m - \rho_c}{\rho_m} \cdot \frac{t_c}{r_0} \cdot (l + 2) = 0.0036 \cdot (l + 2).
\]

We only expect to see the effect of subcrustal mass anomalies for degrees \( l < 10 \) (as discussed in section 2) and this is also approximately the limit of resolution of the tomography model (further discussed below). For \( l = 9 \), the relative error is \( \approx 4\% \), and it becomes smaller for smaller \( l \).

In this way, Eq. 5 is simplified, but we now wish to express \( h_{l_m} \) as the sum of topography \( T_{l_m} \) relative to geoid – the way topography is usually defined, e.g. for the Earth – and geoid \( N_{l_m} \). The geoid in turn can be expressed in terms of gravity potential \( N_{l_m} = -U_{l_m}/g_0 \) whereby \( g_0 = GM/r_0^2 \) is surface gravity and \( U_{l_m} = U_{l_m,1} + U_{l_m,2} \) is the total gravity potential. In this way, Eq. 5 is rewritten as

\[
U_{l_m,2} = \frac{3 \cdot \rho_m}{(2l + 1)\rho} \cdot \left( U_{l_m,1} + U_{l_m,2} - \frac{GM}{r_0^2} \cdot T_{l_m} \right) \quad (6)
\]
Solving this equation for $U_{lm,2}$ gives

$$U_{lm,2} = \frac{\frac{3\rho_m}{(2l+1)\rho} \cdot \left(U_{lm,1} - \frac{GM}{r_0^2} \cdot T_{lm}\right)}{1 - \frac{3\rho_m}{(2l+1)\rho}}$$

and therefore

$$U_{lm} = \frac{U_{lm,1} \cdot \frac{GM}{r_0^2} \cdot \frac{3\rho_m}{(2l+1)\rho} \cdot T_{lm}}{1 - \frac{3\rho_m}{(2l+1)\rho}} =$$

$$= - \frac{GM}{r_0^2} \cdot \frac{3}{(2l+1)\rho} \cdot \frac{T_{lm} \cdot \rho_m + \int_{r_b}^{r_0} \rho_{lm}(r) \cdot \left(\frac{r}{r_0}\right)^{l+2} \cdot dr}{1 - \frac{3\rho_m}{(2l+1)\rho}}$$

The denominator in the last factor is due to so-called self-gravitation. This can be understood because topography is defined relative to the geoid, and the geoid itself is a departure from spherical symmetry. Hence the total geoid is amplified by a factor $> 1$ compared to the equation if topography was defined relative to the spherical shape.

Topography $T_{lm}$ is caused by radial non-hydrostatic stresses $\tau_{r,lm}$ acting on the lithosphere. In the case the elastic strength of the lithosphere can be neglected, the relation between radial stress (at constant depth, i.e. at a constant gravity potential) and topography (relative to the geoid) is simply $T_{0,lm} = \tau_{r,lm}/(\rho_m g_0)$, but topography is reduced for a lithosphere with non-negligible elastic strength. In particular Turcotte et al. (1981) show that for the Moon membrane stresses play an important role in reducing topography. The effect of an elastic lithosphere on topography and hence gravity is also discussed by Zhong (2002) and Golle et al. (2012). This reduction can be described by a “degree of compensation” $f_{el}$ that depends on spherical harmonic degree $l$ and lithosphere elastic thickness $t_e$. In the appendix, we show
how the formalism of Turcotte et al. (1981) can be modified to compute the
degree of compensation for internal loads.

If elastic deformation of the lithosphere occurs relative to a spherical
reference shape, we can write \( T_{lm} + N_{lm} = f_{el} \cdot (T_{0,lm} + N_{lm}) \) and therefore
\( T_{lm} = f_{el} \cdot T_{0,lm} + (f_{el} - 1) \cdot N_{lm} = f_{el} \cdot T_{0,lm} + (1 - f_{el}) \cdot U_{lm}/g_0 \). Inserting
this expression into Eq. 8 and solving for \( U_{lm} \) gives

\[
U_{lm} = -\frac{GM}{r_0^2} \cdot \frac{3}{(2l+1)\bar{\rho}} \cdot \frac{f_{el} \cdot T_{0,lm} \cdot \rho_m + \int_{r_h}^{r_0} \rho_{tm}(r) \cdot \left( \frac{r}{r_0} \right)^{l+2} dr}{1 - f_{el} \cdot \frac{3\rho_m}{(2l+1)\bar{\rho}}}.
\]

Under certain circumstances the relation between \( T_{0,lm} \) and density anoma-
lies \( \rho_{lm} \) can be expressed in terms of “topography kernels” \( K_{t,0,l}(r) \):

\[
T_{0,lm} = \frac{1}{\rho_m} \cdot \int \rho_{lm}(r) \cdot K_{t,0,l}(r) dr.
\]

Similarly, the relation between gravity potential and density anomalies can
be expressed in terms of “geoid kernels” \( K_l(r) \):

\[
U_{lm} = -\frac{GM}{r_0^2} \cdot \frac{3}{(2l+1)\bar{\rho}} \cdot \int \rho_{lm}(r) \cdot K_l(r) dr.
\]

Cases where this kernel formulation is possible include uncompensated den-
sity anomalies, isostatically compensated anomalies and anomalies in a vis-
cous mantle with only radial viscosity variations (Richards and Hager, 1984)
that may be overlain by an elastic lithosphere (Steinberger et al., 2010). For
anomalies isostatically compensated at the surface, topography kernels are
\( K_{t,iso,l}(r) = -(r/r_0)^2 \cdot (g(r)/g_0) \), accounting for smaller surface area at smaller
radius and gravity acceleration \( g(r) \) (see appendix A) decreasing with depth
(they would be 1 for constant gravity in cartesian geometry). For uncompen-
sated anomalies, they are obviously zero. The computation of topography
kernels for a viscous lunar mantle follows the approach of Richards and Hager (1984) but has been modified to account for an elastic lithosphere (Zhong, 2002; Steinberger et al., 2010). More details are given in appendix B. Since it is not clear which (if any) part of the lunar mantle is convecting, we will consider all three cases (no compensation, isostatic compensation, viscous flow beneath elastic lithosphere). Expressing in Eq. 9 both gravity potential and topography in terms of kernels (Eqs. 11 and 10) we can relate geoid kernels to topography kernels

\[ K_t(r) = \frac{f_{el} \cdot K_{01}(r) + \left(\frac{r}{r_0}\right)^{l+2}}{1 - f_{el} \cdot \frac{3 \rho_m}{(2l+1)\bar{\rho}}} \]

(12)

In combination, Eq. 11 and 12 can now be used to compute the geoid, if we know (a) internal density anomalies \(\rho_{lm}\), (b) the degree of compensation \(f_{el}\) for the lithosphere, and (c) topography kernels \(K_{01}\). Geoid kernels are shown in Fig. 6. In the case of uncompensated density anomalies Eq. 12 simplifies to

\[ K_{unc,l}(r) = \left(\frac{r}{r_0}\right)^{l+2} \]

(13)

(red lines in Fig. 6 – positive density anomalies always cause a positive geoid).

In the case of isostatically compensated anomalies Eq. 12 becomes

\[ K_{iso,l}(r) = -\left(\frac{r}{r_0}\right)^2 \cdot \frac{g(r) + \left(\frac{r}{r_0}\right)^{l+2}}{1 - \frac{3 \rho_m}{(2l+1)\bar{\rho}}} = K_{t,iso,l}(r) \cdot \frac{1 - \left(\frac{r}{r_0}\right)^l \cdot \frac{g_0}{g(r)}}{1 - \frac{3 \rho_m}{(2l+1)\bar{\rho}}} \]

(14)

(green lines in Fig. 6 – positive density anomalies always cause a negative geoid). The kernels for a viscous mantle beneath an elastic lithosphere are intermediate between these two cases: The thicker the elastic lithosphere, the closer the kernels are to the case of uncompensated density anomalies.
For small elastic thickness, the negative minimum of the kernels is more pronounced. Kernels are shown for degrees 2, 3, 5 and 9 as the kernels for intermediate degrees are similar and intermediate. Results depend less strongly on thermal thickness $t_l$ (i.e., concerning viscosity structure) and cutoff viscosity of the lithosphere; even if we increase $t_l$ to 1000 km (corresponding to the occurrence of deep moonquakes) or increase cutoff viscosity to $10^{26}$ Pas, resulting kernels look still rather similar. So we use given values of $t_l = 240$ km and cutoff viscosity to $10^{23}$ Pas for all cases (see appendix and Fig. S3 for more details on the radial viscosity structure); the set of cases included in Fig. 6 should appropriately cover the range of kernel shapes that can be expected, regardless of whether the lunar mantle is still convecting, and if so, at what depth.

Hereby the cases of isostatic compensation and the uncompensated case are unrealistic end-member cases. Isostatic topographies are also slightly over-estimated because we assumed that the isostatic compensation is entirely made by the upper surface whereas the deformation should be distributed between the top and the bottom surfaces. However, because of the small core size, isostatic compensation at the CMB should not affect results by much. It will mainly play a role for mass anomalies near the CMB, but these have a small effect on surface topography and geoid anyway. The error made can be estimated from the green curves in Fig. 6: If isostatic compensation at the CMB was properly accounted for, these should reach a value zero at the CMB (bottom of each panel). But since we disregard it, the green curves in Fig. 6 remain slightly above zero.

For our intermediate cases, we first use a viscous rheology to compute ra-
Figure 6: Geoid kernels for uncompensated density anomalies (Eq. 13), isostatic compensation (Eq. 14) and three cases of viscous mantle overlain by lithosphere with elastic thickness $t_e$ (Eq. 12). In these cases, the contributions of internal loads and deflections of the surface are both considered. Surface stresses are computed following the approach of Hager and O’Connell (1981) and Richards and Hager (1984), with radial viscosity structure as in Fig. S2 and gravity acceleration as in Fig. S3, also considering effects of compressibility. Compared to the case without elastic lithosphere, resulting topography is reduced by a factor $f_{el}$. The relation of $f_{el}$ to $t_e$ and $l$ is derived following the approach of Turcotte et al. (1981) which also considers the presence of membrane stresses, and which has been modified to account for the presence of internal rather than external loads on the lithosphere.
dial stresses, and in a second step assume an elastic lithosphere to compute
dynamic topography caused by these stresses. More realistically, the litho-
sphere should have viscoelastic rheology, or be treated as an elastic layer
overlying the viscous mantle, in a single step. However, our approximation
should still be viable: Firstly, among the cases tested (and discussed above)
result show little dependence on lithosphere thermal thickness and viscosity,
so we expect results should remain very similar for a viscous mantle beneath
an elastic lithosphere, at least as long as radial stresses are caused by density
anomalies within the viscous mantle. For density anomalies within the elastic
lithosphere our approach may not be entirely appropriate. However, here we
note that geoid kernels for viscous flow (without any elastic lithosphere; not
shown) and isostatic compensation are similar down to a depth \( \approx 400 \) km for
\( l = 2 \), decreasing to \( \approx 150–200 \) km for \( l = 9 \). Hence we expect that even if
density anomalies are within an elastic lithosphere, resulting topography and
geoid should still remain similar. We also note that in our preferred cases
(see results section) most density anomalies within the elastic lithosphere are
excluded.

Kernels for a viscous mantle and elastic lithosphere were computed with
Young’s modulus \( E = 6.5 \cdot 10^{10} \) Pa and Poisson’s ratio \( \nu = 0.25 \) adopted from
Turcotte et al. (1981). If Young’s modulus is higher, the degree of compensa-
tion is reduced and kernels become more similar to those for uncompensated
density anomalies. Young’s modulus in the lunar lithosphere, and its depth
dependence, is discussed in Pritchard and Stevenson (2000).

The geoid-topography ratio for topography isostatically compensated by
density anomalies in the mantle at depth $z_0 = r_0 - r$ is

$$\text{GTR} = \frac{3 \cdot \rho_m}{(2l + 1) \cdot \bar{\rho}} \cdot \frac{K_{\text{iso},l}(r)}{K_{t,\text{iso},l}(r)} = \frac{1 - \left(1 - \frac{z_0}{r_0}\right)^l \cdot \frac{g_0}{g(r)}}{(2l+1)\bar{\rho} - 1}$$ \hspace{1cm} (15)

In analogy, the geoid-topography ratio for topography due to crustal thickness variations is given by Eq. 3, where we have also neglected the decrease of gravity with depth.

5. Results: Comparison of geoid predictions with observations

In order to assess which degrees to consider in the following, we first compute geoid power spectra based on the Zhao et al. (2012) tomography model. Since at this point we are only interested in how computed power depends on spherical harmonic degree (and not in absolute magnitude), we simply choose a conversion factor $C = 1$. A lower value of $C$ would simply correspond to shifting curves downward. Fig. 7 shows results for three of the cases for which kernels are shown in Fig. 6, and also for the individual layers of the tomography model (converted to density) without upward continuation, i.e. for coefficients

$$C^{G}_{lm} = \frac{3}{(2l + 1)\bar{\rho}} \cdot \delta \rho_{lm,i} \cdot \Delta r_i$$ \hspace{1cm} (16)

whereby $\delta \rho_{lm,i}$ are expansion coefficients of the density anomalies inferred from the tomography model in layer $i$ and $\Delta r_i$ is layer thickness, and $S^{G}_{lm}$ in analogy. In contrast to the observed power spectrum, which becomes rather flat above degree 10, the power predicted from the tomography model continues to drop with increasing degree. Given the limited resolution of the
tomography model, this is not surprising. Hence we do not expect reliable results for degree $\approx 10$ and higher.

Given the resolution of the tomography models and the degree range where we think, based on Figs. 2 and 7, that a deeper than crustal origin of geoid undulations is possible, we now limit our analysis to degrees $l \leq 9$. The top row in Fig. 8 shows observation-based topography and geoid, filtered to these low degrees, whereas in the middle row, we show examples for modelled topography and geoid. In the case of a rigid lithosphere where only uncompensated internal density anomalies contribute to the geoid (part F), negative density anomalies always cause negative geoid and vice versa.

In the bottom row, correlation and ratio of predicted and observed geoid are shown for a larger number of cases. We consider the limiting cases of uncompensated density anomalies and isostatic compensation (where negative density anomalies always correspond to positive geoid and vice versa, because the effect of isostatic topography on the geoid is always dominant).

We also consider the intermediate cases with a viscous mantle and elastic lithosphere. With increasing elastic thickness, these cases approach the “uncompensated” limit. In addition, we consider cases where density anomalies are isostatically compensated above a certain depth and uncompensated below. The scenario that would approximately justify such an assumption is that shallow density anomalies formed during early evolution of the Moon, when its lithosphere was still thin such that they could be partly isostatically compensated, and later on got frozen in. In contrast, if convection continued below a thickening lithosphere, associated deeper anomalies would deform the lithosphere much less and be nearly uncompensated. We compare our
Figure 7: Square-root of geoid power computed from the Zhao et al. (2012) tomography model assuming a conversion factor $C = 1$, or the model based on moonquake locations below 225 km depth (red dotted line only). Green, red and violet lines are for the same cases as the geoid kernels in Fig. 6. For the red and violet lines, we either convert all anomalies to density anomalies (upper, continuous lines) or only those below 225 km depth (lower, dashed lines). The green line is for all anomalies. Grey, black and blue lines are for individual layers without compensation and without upward continuation. For comparison, the observed spectrum after mascons have been removed is shown as brown continuous line (same as red line in Fig. 2). The brown dotted line between $l = 2$ and 3 is obtained after flattening has also been removed (same as orange line in Fig. 2).
computations with the observed geoid for three cases: In the first case (part G), we keep all coefficients, in the second case (part H) we reduce $C_{20}^G$ to one third of its value, in the third case (parts J and K) we set it to zero. The third case corresponds to assuming that the flattening is a "fossil bulge", the second case that this is partly so.

We find the highest correlation between modelled and observed geoid (0.51) in the case (shown in Fig. 8 F, and indicated by black boxes in parts J and K) where flattening has been removed and anomalies are uncompensated and only anomalies below 225 km depth are considered, the second highest (0.46) in the similar case (shown in Fig. 8 D and E, and indicated by grey boxes in parts J and K) with a thick elastic lithosphere ($t_e = 240$ km). If we increase the Young’s modulus from $6.5 \times 10^{10}$ Pas (Turcotte et al., 1981) to $1.6 \times 10^{11}$ Pas, correlation in the second case somewhat increases to 0.54.

However, in these cases, predicted geoid amplitudes are lower than observed, although we use a conversion factor $C = 1$ which is rather high, at least for thermal anomalies. Using a lower conversion factor simply corresponds to reducing amplitude, or changing the color bars for geoid and ratio in Fig. 8 accordingly. The predicted geoid for these two cases is shown in parts E and F. The positive correlation corresponds to some similarities in the patterns of the model predictions and the corresponding observation (part C). The spectra for these two best-fit cases (lower red and violet lines in Fig. 7) again show that predicted amplitudes are too low, however the shape of the spectrum approximately follows the observed one up to $l \approx 7$. For higher degrees, power drops more strongly than observed, probably due to limited tomography resolution. Predicted topography for the second case
is shown in part D, while in the first case, zero topography is assumed. The predicted topography has much smaller amplitude than observed topography, and the predicted pattern is similar to opposite to the observed one. This would mean that topography has mostly shallow origin, and is not caused by mantle density anomalies.

Given that – at least for the Earth’s upper mantle a conversion factor $C$ of around 0.22 is estimated (Steinberger and Calderwood, 2006), predicted geoid amplitudes are more realistic if anomalies within the lithosphere are at least partly kept. However, this comes at the price of reducing correlation. In contrast, if we compare to the geoid with the flattening term included, we find negative or near-zero correlations in case of uncompensated anomalies or thick elastic lithosphere. For thin elastic lithosphere or isostatic compensation, though, correlation becomes again positive, but only reaches values up to $\approx 0.2$. But assuming thin lithosphere or isostatic compensation, positive correlations are only possible if the flattening term is included; if it is removed, correlation becomes negative.

Interestingly, we find that we can obtain even higher correlations with a geoid model, which is based on our tentative density model derived from moonquake locations only. Fig. 9 (a) shows that observed moonquakes are clustered in a region centered on the center of the near side. However, it is not clear whether the region near its antipode is really nearly aseismic, or whether just moonquakes in that region could not be observed with the available Apollo seismometers (Nakamura, 2005). Assuming uncompensated density anomalies, this density model yields a predicted geoid high also near the center of the near side (Fig. 9 b) very near the actual nearside maximum.
Figure 8: Geoid and topography up to spherical harmonic degree l=9. In all maps, Procellarum KREEP terrane is outlined in black, mare units are outlined in white. A and B: Observed topography and geoid after removal of mascons (same as Fig. 1 C and D, but filtered to only retain long wavelength). C: Also the flattening term $C_{20}^G$ has been removed. D and E: Modelled topography and geoid with elastic lithosphere thickness $t_e = 240$ km and only anomalies below 225 km depth considered with conversion factor $C = 1$. F: Modelled geoid with uncompensated density anomalies; other assumptions as in E. Bottom row: Correlation (G to J) and ratio (K) of predicted and observed geoid for the five cases in Fig. 6 and $C = 1$ on the near side. G: $C_{20}^G$ included; H: $C_{20}^G$ multiplied with $1/3$; J and K: $C_{20}^G$ set to zero. Four rows are depths above which either isostatic compensation is assumed with no compensation below (columns marked iso-unc) or above which we set $C = 0$ (other columns). The grey boxes mark the case shown in part D and E, the black boxes the case in part F.
of the residual geoid (Fig. 8 B). Since the geoid by definition does not have a degree-one term, there is also a compensating far-side geoid high, again approximately corresponding to the observed one. The correlation with the actual geoid (flattening included) is 0.87 on the near side and 0.72 over the entire surface. Even if the flattening term $C_{20}^G$ of the actual geoid is reduced to one third, near-side correlation is still 0.6. However, most moonquakes below 225 km occur at depths below 800 km where geoid kernels are very small, particularly for higher degrees. Therefore, a density anomaly 3.65% has to be assumed in order to match the observed geoid amplitude. This is probably unrealistically large, corresponding to $\sim 1000$ K in case of a thermal anomaly. The predicted geoid is strongly dominated by degree two. This is also evident from the dotted line (representing the density model based on moonquake locations) in Fig. 7: The geoid power spectrum for that model decreases much more strongly with $l$ than the observed spectrum. The comparatively high correlations can result, because the observed geoid also has a large degree-two component.

Fig. 2 shows that, for degrees 3-9, after the effect of mascons has been removed, geoid and topography are highly correlated, and that after subtracting the effect of topography, assumed to be isostatically compensated at depth 70 km, geoid power becomes substantially less in the same degree range. We hence also do the same analysis as in Fig. 8 for the residual geoid, assuming isostatic compensation either at depth 70 km (as in Fig. 2), or at depth 50 km, the crustal thickness from Wieczorek et al. (2006)). However, we find in both cases generally a worse fit than in Fig. 8. Interestingly, the best correlation with the residual geoid (0.27 for compensation depth 70 km,
Figure 9: (a) Distribution of moonquake epicenters (Nakamura, 2005). (b) Predicted geoid inferred from an uncompensated density model based on the moonquake epicenters only. A density anomaly of 3.65% is assumed nearby the moonquakes (see section 2 for details), whereas density anomalies in the uppermost 225 km are removed. Procellarum KREEP terrane is outlined in black.

0.26 for 50 km) is obtained for the case which is quite the opposite to the best-fit case in Fig. 8 – isostatic compensation, tomography at all depths included, observed flattening included. In this case, using $C = 1$, predicted topography is of similar magnitude, but unrelated to observed topography, making this model less plausible.

6. Discussion

One of the basic assumptions that underlies our analysis is that we convert seismic velocity to density anomalies. This assumption is reasonable, if both are caused by temperature anomalies, but becomes questionable if both thermal and compositional density anomalies play a role. For example, it is regarded a reasonable assumption for the large part of the Earth mantle, where thermal density anomalies are thought to be dominant. But,
in contrast, there are positive large seismic velocity anomalies in the Earth’s continental lithosphere probably without corresponding density anomalies (Jordan, 1988). And in the Large Low Shear wave Velocity Provinces of the lowermost mantle, negative seismic velocity anomalies are likely even associated with positive density anomalies (Ishii and Tromp, 2004), i.e. these are likely even more dense than the surrounding mantle. Also for the lunar mantle, compositional anomalies have been suggested as a cause for seismic heterogeneities (Sakamaki et al., 2010).

Therefore, converting velocity to density anomalies should be regarded as an assumption, that is not necessarily true, but at least reasonable, and we are testing here whether and under what circumstances it leads to reasonable model predictions. Further, in order to account for the possibility that – as is presumed for the Earth – most seismic velocity anomalies below a certain depth correspond to density anomalies, whereas at shallower depth there are large anomalies without corresponding density anomalies (Jordan, 1988), we also consider cases where we only convert seismic velocity anomalies to density anomalies below a given depth, and disregard them above.

We find the highest correlations between predicted and observed geoid for the case that there is either a thick elastic lithosphere \((t_e = 240 \text{ km})\), or density anomalies are uncompensated, corresponding to even thicker lithosphere. This appears reasonable, given that also other observations indicate a rather thick lunar lithosphere. The occurrence of moonquakes deep inside the moon (Fig. 9), though, may be due to alternative mechanisms that allow for brittle failure in an otherwise ductile environment (Frohlich and Nakamura, 2009). The fact that correlations are higher if density anomalies
above 225 km are excluded could mean that, similar to the Earth, above 225 km, compositional density anomalies play a larger role, such that a simple velocity - density conversion is less appropriate there. Correlation in this case is higher if the flattening term of the observed geoid is removed. This would point towards flattening being a “fossil” remains from an earlier time (Lambeck and Pullan, 1980), which subsequently was “frozen in” due to a lithosphere that had gradually thickened.

We can obtain an even higher correlation with the observed geoid, if we base our density model on the moonquake distribution only, assigning higher densities to volumes around the hypocenters of deep moonquakes (> 225 km depth). However, in this case, matching geoid amplitudes requires unrealistically high density anomalies, because most of these moonquakes and the inferred density anomalies occur at great depth. Density anomalies could have a more realistic magnitude, if they extend to shallower depths above the deep moonquake hypocenters. In contrast to the tomography-based model, we find the highest correlation here if the flattening term is included, meaning that the geoid flattening would be due to internal density anomalies, different from the results based on the tomography model.

Also in the case of the density model based on tomography, the predicted geoid amplitude is much too low. This could mean that the tomography model is strongly “damped” with amplitudes much lower than in reality – an effect that is known to affect tomography models on Earth. The tomography model has amplitudes of the order 1% and given conversion factors considered appropriate for the Earth’s upper mantle this corresponds to temperature anomalies of only ≈ 70-100 K, i.e. there could be some damping if actual
anomalies are higher. But it could also mean that mantle density anomalies
beneath ≈ 225 km only contribute a small part to the geoid, and it mostly
originates at shallower depth. Given the data available, we cannot resolve
this issue.

For the preferred cases (framed in grey and black in the bottom row, and
also shown in the middle row in Fig. 8) negative geoid corresponds to dom-
inantly negative density anomalies for spherical harmonic degrees two and
higher (see Fig. 6) which – if there is viscous flow – would correspond to up-
wellings and upward deflection of the lithosphere, and vice versa, unless there
is a strong degree-one component in density anomalies (i.e. positive in one
hemisphere, negative in the other) which, by definition, will not have a geoid
signature. In the preferred cases such a negative geoid anomaly is predicted
centered around 50° W, 10° N, and in the case of a thick elastic lithosphere,
also an upward deflection of the lithosphere is predicted there. The actual
geoid (with flattening term removed) has a minimum further west, around
70° W. The observed minimum is near the western edge of the Procellarum
KREEP terrane, the modelled minimum (and center of upward lithosphere
deflection) is closer to its center. Thus the latter could correspond to the
positive thermal anomaly proposed to underlie the KREEP terrane (Wiecz-
zorek and Phillips, 2000). The predicted geoid highs to the east would then
correspond to positive density anomalies. In the case of a viscous mantle,
this would correspond to downward flow to compensate for the upward flow
further west (Fig. S4). However, the largest predicted geoid maximum is still
inside the KREEP terrane, and the observed maximum is even closer to its
center.
Such a positive geoid anomaly, which, for our preferred models, would correspond to positive density anomalies, and presumably cold and possibly sinking material, contrasts with the suggestion that the KREEP terrane is underlain by hotter-than-average material. However, we have to consider the possibility of additional degree-one density anomalies and flow, with hotter and possibly upwelling material on the near side (Laneuville et al., 2013) which would not be visible in the geoid and could not be detected from the available seismic data, since we have no information about the far side.

In the northern half of the near side, we also find lower-than-average geoid-topography correlation and higher-than-average ratio (Fig. 4). If both geoid and topography are largely due to crustal thickness variations and other isostatically compensated lithospheric density variations, we expect a high correlation and the higher geoid-topography ratio the deeper the compensation level. The regionally low correlation and high ratio could be caused, if in that region – despite our attempt to eliminate the effect of mascons – low topography is still isostatically over-compensated: If the effect of mascons is included, the strong over-compensation even leads to a regionally negative correlation and much higher ratio, in particular in the degree range where mascons have most power. This degree range (centered on 10-11) can be estimated from the range where in Fig. 2 the blue curves (mascons included) and red curves (mascons removed) are most different. But the low correlation and high ratio could also be an indication that in this region, there are stronger-than-average mantle anomalies, perhaps more dominated by negative, hot anomalies and (past or still ongoing) upwelling in the western part beneath the KREEP terrane and positive, cold anomalies further to the east.
(see also Fig. 5 and Fig.S4). Whether or not convection is still ongoing, or has stopped or at least greatly slowed down but with anomalies still remaining within a mostly rigid mantle cannot be decided from our analysis: In the preferred cases, the predicted geoid is largely or fully due to internal density heterogeneities, and at most to a small part due to boundary deflections, so it makes little or no difference, whether, and if so, in which regions, the lunar mantle is still convecting.

Given that a large part of the geoid, at least in the degree range 3-9, can also be explained by shallow isostatic compensation, we expected that perhaps, if the effect of isostatically compensated topography is removed, the correlation of our geoid predictions with the remaining “residual” geoid is even higher. However, we found that generally the fit gets much worse. We think that this failure to obtain an improved fit could be due to a combination of two causes. Firstly, degree two, with its rather strong power, is not well explained by shallow compensation. Secondly, in the whole degree range 2-9 around the region where the seismic stations were deployed (Fig. 4), geoid and topography are less well correlated, hence isostatic compensation probably explains the geoid less well.

More specifically, in this region, with abundant lunar maria and low topography (Fig. 1), assuming isostatic compensation yields a residual geoid that is more strongly positive than the actual geoid. However, the lunar mascons are regions of low topography with strongly positive geoid, implying isostatic “over-compensation”, and if a similar but smaller effect is more common in that region, trying to separate off the mantle contribution of the geoid by assuming isostatically compensated topography may be inappropri-
ate, at least in this region where seismic coverage is best. Moreover, the issue
is further complicated by variable crustal densities. Wieczorek et al. (2013)
find densities for the highland crust much lower than previous crustal density
estimates. Also, it would be inappropriate for our purpose to subtract a grav-
ity contribution computed from crustal density and thickness models, since
these are not independent but in turn derived from gravity and topography.

7. Conclusions and Outlook

We have investigated here the question of how much of the lunar geoid
and possibly topography has a “deep” (meaning substantially deeper than
crustal levels) origin. The investigation was motivated by the observations
that on one hand, geoid and topography are highly correlated, and a large
fraction of the geoid be explained by isostatically compensated topography,
in particular up until spherical harmonic degree $l = 9$. On the other hand,
the geoid power spectrum for low degrees has a different slope than at higher
degrees. This could possibly indicate a dominantly deep origin up until
spherical harmonic degree $l \approx 9$.

We address this question by comparing observed geoid and topography
with predictions based on a tomography model. The “observed” geoid is
modified by subtracting the effect of “mascons” which are almost certainly
shallow features, as across them geoid and topography are clearly related.
We also optionally removed fully or partly the degree two order zero term,
which could be a “fossil” feature caused early in the Moon’s history. We
assume for simplicity a linear relation between relative seismic velocity and
relative density variations. But we also consider that this linear relation only
holds beneath a certain depth, disregarding seismic anomalies above. We
calculate the geoid for a number of assumptions – uncompensated density
anomalies, isostatic compensation, or a combination of both, or intermediate
cases with an elastic lithosphere of various thickness above a viscous mantle.

We find the highest correlation if we assume uncompensated density
anomalies or a thick elastic lithosphere, if we do not consider shallow seismic
anomalies, and if we compare with the geoid where the degree two order zero
term has been set to zero. That the highest correlation occurs for assuming a
very thick lithosphere is not surprising, given that also other evidence points
to a thick lunar lithosphere. Also, the fact that including shallow anomali-
ies worsens correlation can be readily explained if – as is also presumed for
the Earth – shallow seismic anomalies are due to thermal and compositional
anomalies, and the latter are more prevalent at shallow depth. Our preferred
model is consistent with the idea that hotter-than average mantle underlies
the western part of the Procellarum KREEP region in the northwest of the
lunar near side, where lunar maria are abundant. In this model, positive dy-
namic topography is predicted where actual topography is below the mean,
but has an amplitude of no more than a few hundred meters. To the east of
this region, our model features positive density anomalies overlain by positive
geoid. These could correspond to the downgoing limb of a convection cell,
with the main upwelling further west. Geoid-topography correlation lower
than average and ratio higher around the northern part of the lunar near
side could indicate that beneath this region, density anomalies are stronger
than elsewhere in the deep lunar mantle.

However, we have to be careful not to over-interpret our results, as the
geoid amplitudes of our preferred model cases are much lower than observed. This could either indicate that the tomography model is strongly damped, with amplitudes much lower than in reality, or that a large part of the geoid has a shallow origin, due to topography isostatically compensated for example due to crustal thickness variations. So we cannot yet present any definite conclusions regarding the depth of origin of the long-wavelength \((l \leq 9)\) lunar gravity field. We are perhaps now in a similar situation for the Moon as were in the 1970s for the Earth’s mantle, when the principal large-scale features of mantle anomalies were only beginning to become apparent – the first successful predictions of large-scale geoid anomalies due to mantle density structure were only presented in the 1980s. We anticipate that this failure, and the promise of obtaining more significant results with better data should serve as a motivation to undertake more efforts to collect such data, not only on the Moon but also on other planets. One such effort is the InSight mission to Mars scheduled in 2016, and we hope our paper can illustrate a way how information gathered through such programs can be used for learning more about planetary interiors.

Acknowledgements

Figures were made using GMT (Wessel and Smith, 1998). B.S. and S.C.W. are supported by the Norwegian Research Council through a Center of Excellence grant to The Centre for Earth Evolution and Dynamics and partially by the European Research Council under the European Union’s Seventh Framework Program (FP7/2007-2013)/ERC Grant agreement 267631 (Beyond Plate Tectonics). D.Z. is supported by a grant (Kiban-S 11050123)
from Japan Society for the Promotion of Science (JSPS) and a grant (Shin-
Gakujutsu 26106005) from MEXT. We thank three anonymous reviewers for
constructive and detailed comments leading to considerable improvements of
the manuscript.

Appendix A. A model for the radial structure of the Moon

We follow here a strategy that we have previously in a similar fashion
applied to Venus and Mars (Steinberger et al., 2010): We assume a radial
mantle viscosity profile $\eta(r) \sim \exp(rH/(RT))$ where $H$ is activation enthalpy,
$R$ is the universal gas constant, $T$ is temperature and $r$ is a constant for
which we use here a value $1/3.5$. In the case of a non-linear stress-strain
relationship, this is an “effective” viscosity (Christensen, 1983). The pressure
range in the mantle of the Moon corresponds to the Earth’s upper mantle,
for which often a dislocation creep mechanism is assumed and $r = 1/3.5$
should be approximately appropriate for effective viscosity. We compute
viscosity based on a temperature profile that is adiabatic in the interior and
with thermal boundary layers, and assume that the pressure dependence of
adiabatic temperature and activation enthalpy (Fig. S1), and the pressure
and temperature dependence of viscosity are the same as derived for the
Earth (Steinberger and Calderwood, 2006; Calderwood, 1999).

Mantle temperature and density (Fig. S1) as well as thermal expansivity
as a function of pressure are obtained from a self-consistent model (Schmeling
et al., 2003; Steinberger and Calderwood, 2006) based on available mineral
physics data. Core density as a function of pressure is extrapolated from the
relation inferred for the Earth’s core based on PREM (Dziewonski and An-
Figure S1: Adiabatic temperature (purple), activation enthalpy (red) and density (green) as a function of pressure.
Pressure, gravity and density are then downward-integrated for the given pressure-density relation, and a given crustal thickness 50 km and density 2900 kg/m$^3$. This is similar to Table 3.10 of Wieczorek et al. (2006). We adjust mantle and core density such as to match the known moment of inertia factor 0.3932 (Konopliv et al., 1998), seismically determined core radius 330 km (Weber et al., 2011) and total mass. This yields the depth profiles of pressure, gravity and density shown in Fig. S2. For simplicity, we do not distinguish between outer and inner core. For the Moon, the core densities obtained in that way ($6.41 - 6.43 \cdot 10^3$ kg/m$^3$) are similar to the average core density found by Weber et al. (2011). Uppermost mantle density is $3.31 \cdot 10^3$kg/m$^3$, which is very similar to Wieczorek et al. (2006).

Resulting profiles of temperature and viscosity as a function of depth are shown in Fig. S3. No thermal boundary layer at the bottom of the mantle is assumed, given the small core radius. We note that the inferred thermal structure may possibly correspond to a “fossil” one, in the case that convection has stopped by now. In this way, the thin elastic and thermal thickness may correspond to earlier times. The assumed elastic thicknesses (Table 1) are similar to and somewhat larger than the two elastic thickness estimates given by e.g. Freed et al. (2001) for two times earlier in the Moon’s history. However, a detailed treatment of structural relaxation should take secular cooling into account (Kamata et al., 2012).
Figure S2: Models for density, gravity and pressure as a function of depth for the Moon.
Figure S3: Inferred temperature and viscosity profiles for the mantle of the Moon. The straight lines are for adiabatic temperature, the curved lines with thermal boundary layers. The green line is the approximation with layers of constant viscosity actually used in the flow computation. Lithosphere viscosity is cut off at $10^{23}$ Pas, however resulting geoid kernels remain very similar even for much higher cutoff viscosities.
Appendix B. Computation of flow and topography for a viscous lunar mantle overlain by an elastic lithosphere

The traditional viscous flow modelling approach (Hager and O'Connell, 1981; Richards and Hager, 1984) uses zero normal displacement as surface boundary condition, which implies normal stresses at the surface. In the case of a rigid lid, which is appropriate for the Moon, the other surface boundary condition is zero tangential flow. However, the normal stresses are interpreted as representing surface topography, and the contribution of this surface topography to the geoid is also considered.

The effect of density anomalies at a given depth (radius \( r \)) and spherical harmonic degree \( l \) on topography can then be expressed in terms of topography kernels \( K_{0,l}(r) \) (Eq. 10). These topography kernels can be computed from models of viscous mantle flow for given viscosity profiles. They only depend on relative variation of viscosity with depth, not on the absolute viscosity values, but flow speeds are proportional to these. Fig. S4 shows a cross section through a density and flow model corresponding to the cases shown in Fig. 8 D to F. Here we also consider the effect that an elastic lithosphere combined with a viscous mantle has on geoid kernels – a non-standard formulation that was first published and demonstrated effective by Zhong (2002). In this case, the surface deflection is reduced compared to a purely viscous mantle.

The effect of membrane stresses (Turcotte et al., 1981) is also considered, which causes that topography near the surface is substantially less than in the case without elastic lithosphere, even for the lowest degrees. Only a small fraction of surface topography on the Moon is thus isostatically compensated
Figure S4: Vertical cross section at latitude 8.3° N through the density and flow field for the cases in Fig. 8 D to F. Arrow length 10 degrees of arc corresponds to 1 cm/yr.
(Zhong and Zuber, 2000). However, we expect a larger degree of compensa-
tion if the stresses act from inside the lithosphere: Turcotte et al. (1981) state
that they implicitly assume that the region between zero level and downward
displacement of the lithosphere is filled with crust of density $\rho_c$. For internal
loads, it appears more appropriate to not assume such a fill-in and hence re-
place $\rho_m - \rho_c$ by $\rho_m$ in their Eq. (3). Accordingly, we compute the degree of
compensation from their Eq. (27) with $\sigma$ and $\tau$ defined similarly as in their
Eqs. (6) and (7) but with $\rho_m - \rho_c$ replaced by $\rho_m$. Steinberger et al. (2010)
used this modified approach for Mars, but due to a mixup between the two
approaches, the elastic lithosphere thickness for Mars had been incorrectly
given as 102 km, whereas in fact it was 208 km for the results shown in that
paper.

References

Araki, H., et al., 2009. Lunar global shape and polar topography derived

Braden, S.E., Stopar, J.D., Robinson, M.S., Lawrence, S.J., van der Bogert,
within the past 100 million years. Nat. Geosci. 7, 787791.

and composition of the Earth’s mantle. PhD thesis, University of British
Columbia, Dept. of Earth and Ocean Sciences, Vancouver, B.C., Canada.

versus power-law rheology. Earth Planet. Sci. Lett. 64, 153–162.


