INTRODUCTION

This supplementary material provides a detailed explanation of the forward model used in this paper. It aims to describe the physics used within the model. It also expands the description of the parameters from that within the paper, especially the starting parameters. The reason for the choice of these is outlined, as well as the effect this has on the model as it runs forward in time. The model is a one dimensional forward finite difference code written in Matlab which calculates the conductive heat flow through a column of the continental crust, mantle lithosphere and upper most asthenosphere. In principle it is similar to the plate model used to calculate the bathymetry and heat flow of oceanic crust away from a mid ocean ridge (Parsons and Sclater, 1977; Stein and Stein, 1992). However, the starting conditions differ in our model because they describe a continental crust with a plume head placed beneath it.

CALCULATION OF THE HEAT FLOW AND THE SUBSIDENCE

The conductive heat flow through time for the column of rock is calculated from the initial conditions using Fourier’s law combined with the conservation of energy.

\[ \rho C_p \frac{dT}{dt} = \delta \left( \frac{\delta T}{\delta z} \right) + A \]  

From equation (1) the temperature \( T \) of a point in the model at a given depth \( z \) and time \( t \) can be calculated with the thermal conductivity \( k \), the density \( \rho \), the heat capacity \( C_p \) and the radioactive heat production \( A \). In the model the column is discretised using a 1-km resolution. The time stepping in the model is carried out using the Forward Euler method. The new temperature for the \( i \)th point in the model can be calculated using equation (2).

\[ T_{new}^{(i)} = T_{old}^{(i)} + \frac{\Delta t}{\rho^{(i)} C_p^{(i)}} \left[ k_{(i+1/2)} \left( T_{old}^{(i+1)} - T_{old}^{(i)} \right) - k_{(i-1/2)} \left( T_{old}^{(i)} - T_{old}^{(i-1)} \right) \right] + \frac{\Delta t A_{(i)}}{\rho^{(i)} C_p^{(i)}} \]  

The parameters for each point in the model (\( k, \rho, C_p \) and \( A \)) all vary with rock type. The density also varies with temperature as shown in equation (3).

\[ \rho = \rho_0 \left( 1 - \alpha (T - T_0) \right) \]  

(3)
where $\rho_0$ is the reference density at a reference temperature $T_0$ and $\alpha$ is the thermal expansivity.

The elevation of the column from the model can be calculated isostatically by comparing it to a column of mid ocean ridge material using a global average water depth of ~2.7 km overlying a 7 km thick crust composed of basalt, gabbro and mafic cumulates and is directly underlain by depleted upper mantle. This process of using isostacy is similar to that used to calculate the bathymetry away from a mid ocean ridge. When the elevation of the model drops below sea-level a basin is considered to be created. This can be filled with either sediments or water. The fill is not considered when calculating the heat flow, but is used to calculate the subsidence.

BOUNDARY AND INITIAL CONDITIONS, AND PARAMETERS USED

The initial temperature conditions and model depth play an important role in the subsidence patterns produced and therefore must be chosen carefully. As in the plate model the thickness of the model was chosen so that the final lithospheric thickness would match the present day measurements of lithospheric thickness (Artemieva and Mooney, 2001; Priestley and McKenzie, 2006). Lithospheric thickness varies as described in the main paper so an intermediate value of 150 km was used for the present day lithospheric thickness. The 1200 °C isotherm was used as a proxy for the lithospheric thickness (Stein and Stein, 1992). For this isotherm to be positioned at 150 km once the model has stopped cooling a 170 km thick model is used. Essential boundary conditions are used in the model; the temperature at the surface is 0 °C and the temperature at the base of the model is 1381.3 °C which is derived from an adiabatic gradient of 0.3 °C/km and a potential temperature of 1330 °C at the surface. The initial temperature profile has a uniformly high temperature plume head which can be varied in temperature, thickness and depth. The top of the plume head is defined to coincide with the base of the lithosphere. Therefore, a linear temperature gradient is used above the plume from 0 °C at the surface to 1200 °C the point overlying the plume. The temperature profile below the plume follows the adiabatic gradient. An example is shown in Figure 3a in the accompanying paper.

These initial conditions are a simplification of the real situation. The plume head was not emplaced instantaneously and may have cooled as it has thinned the lithosphere and melt has been removed so may not be a uniform temperature. No radioactive heat production is included in the initial crustal geotherm. However, the geotherm reaches a more realistic scenario within the first 0.5 Myrs as the model evolves so this little effect on the subsidence curves produced by the model (Fig. DR1). The dependency of the model results in the initial conditions used is what allows their effects on the subsidence produced to be investigated. Care was taken to make sure that parameters such as the crustal thickness matched the observed data for these (Vyssotski et al., 2006). The present day sedimentary thickness was subtracted from the Moho thickness to give an estimate of the initial crustal thickness. The initial uplift produced by our model of 1500-2000 m matches the ~2000 m shown for the model of a plume head rising beneath juvenile crust neighboring a craton in Burov et al., (2007).Their model includes dynamic uplift so the
agreement shows it does not alter the results from our models significantly. Their models only cover the first 15 Myrs so do not consider the subsidence as the plume cools.

Table DR1. The various parameters used to describe the rock properties.

<table>
<thead>
<tr>
<th></th>
<th>Reference density ($\rho_0$) kg m$^{-3}$</th>
<th>Radioactive heat production (A) W kg$^{-1}$</th>
<th>Coefficient of thermal expansion ($\alpha$) K$^{-1}$</th>
<th>Specific heat ($C_p$) J kg$^{-1}$ K$^{-1}$</th>
<th>Coefficient of conductivity (k) W m$^{-1}$ K$^{-1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sediments</td>
<td>2200</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Water</td>
<td>1030</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Basalt</td>
<td>2900</td>
<td>0.6 x 10$^{-6}$</td>
<td>2.4 x 10$^{-5}$</td>
<td>790</td>
<td>3.1</td>
</tr>
<tr>
<td>Upper crust</td>
<td>2700</td>
<td>1.31 x 10$^{-6}$</td>
<td>2.4 x 10$^{-5}$</td>
<td>790</td>
<td>3.1</td>
</tr>
<tr>
<td>Lower crust</td>
<td>2900</td>
<td>0.6 x 10$^{-6}$</td>
<td>1.6 x 10$^{-5}$</td>
<td>790</td>
<td>2.1</td>
</tr>
<tr>
<td>Mantle peridotite (including plume)</td>
<td>3300</td>
<td>0.006 x 10$^{-6}$</td>
<td>3.3 x 10$^{-5}$</td>
<td>790</td>
<td>3.3</td>
</tr>
</tbody>
</table>

Note: Values for density are taken from Allen and Allen, (2005), the values for heat production, $\alpha$ and $C_p$ and k are from Shaw et al., (1986) and Turcotte and Schubert, (2002). They are similar to the values used for these parameters in Burov et al., (2007).

**MODEL LIMITATIONS**

Due to the nature of the model it does not include the dynamic pressure contribution of the moving plume to the topography. This will have an immediate effect as the plume head impacts the base of the lithosphere contributing to the initial uplift, but will only occur as long as there is active upwelling in the mantle beneath the basin. However, there is no evidence for a plume tail resulting in volcanism after the flood basalts. It has been suggested this may be because the tail moved beneath the pole (Burke and Torsvik, 2004) and its trace is obscured by ice or the plume head may not have given rise to a tail (Burov et al., 2007). Either way dynamic uplift will cease to affect the basin once the upwelling has ceased or has moved. The heat flow and subsidence are only calculated in 1D so the effect of flexure of the crust is ignored. It also assumes that the dominant heat flow is vertical and that there are not significant horizontal temperature gradients. For a basin on the scale of the West Siberian Basin flexure is unlikely to influence the results of the model except close to the margins of the basin, because its size far exceeds the effective elastic thickness of the continental crust. This limitation is similar to the backstripping of a 1D well, and the effects of flexure were shown to negligible by Saunders et al., (2005).
Figure DR1. The evolution of the geotherm with time for the forward model. This demonstrates that the simplified initial conditions quickly evolve into a more realistic temperature profile through the crust and plume head. It also shows the majority of the cooling occurs during the first 200 Myrs.


