Journal of Geophysical Research

Supporting Information for

Lateral variations in thermochemical structure of the Eastern Canadian Shield

Thomas Eeken[1], Saskia Goes[1], Laura Petrescu[2,3], Isabella Altoe[1]

[1]Department of Earth Science and Engineering, Imperial College London, UK
[2]Istituto Nazionale di Geofisica e Vulcanologia, Bologna, Italy , [3]currently at: National Institute for Earth Physics, Bucharest, Romania

Contents of this file

S1. Cluster Analysis

We use the implementation of the k-means method in Matlab. This clustering method iteratively groups the dispersion curves into several sets based on the average distance of a curve to the average of the set. Results depend on the randomly chosen starting estimate of the cluster centres. We therefore run the algorithm a large number of times with different starting centroids and chose the solution with the lowest total within-cluster sum of squared differences as our preferred solution. It also needs to be determined what the optimum number of groups is. When split into two clusters, we find that the dispersion curves are separated into those from the Grenville and those from the Superior. Three clusters lead to a subdivision of the Superior curves into 2 groups. Four groups also split up the Grenville curves, leading to the clusters shown in Figure 3. Table S1 shows the average distance of the curves within a group to the group average. More groups result in average distances to multiple groups that are smaller than the errors on individual dispersion curves (between .02 and .06 depending on period), hence four is our preferred number of clusters. Fig. S1 shows silhouette plots [Rousseeuw, 1987] for our preferred set of clusters (Figure 3).

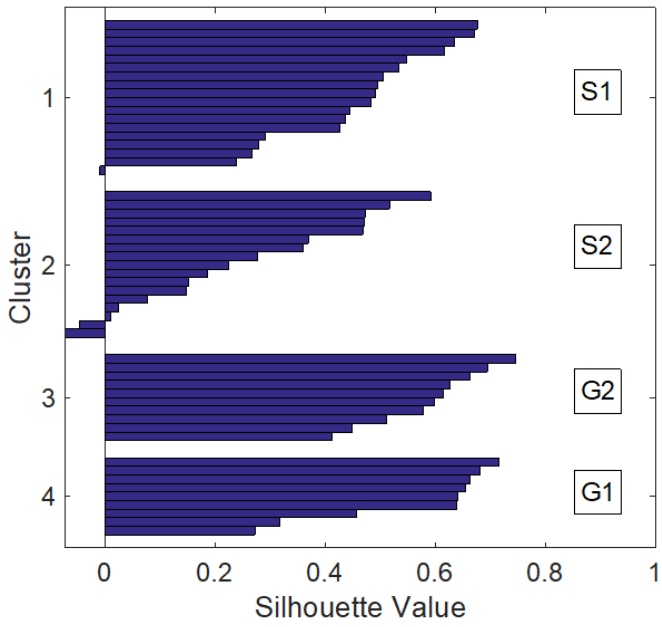

Figure S1: *Silhouette plot that illustrates to what extent the dispersion curves within each cluster resemble the cluster mean. Values lie between +1 (excellent correspondence), and -1 (curves poorly matched to the group), while a value near 0 indicates that the curve may fit just as well in one of the other clusters. If we look at the groups then we can see that G1 and G2 are clearly distinct groups with the curves most likely put in the correct group, S1 is mostly the same but with some poorer fits. S2 is different in that the curves do not all neatly fit to the group average. This may be more of a rest group that Ideally would be split up further, although that would decrease the difference between groups to less than the error on the dispersion curves.*

GROUP\AV. DISTANCE TO GROUP	GROUP S1	GROUP S2	GROUP G1	GROUP G2
Group S1	0.0210	0.0679	0.1399	0.0943
Group S2	0.0796	0.0327	0.2142	0.2259
Group G1	0.1358	0.1984	0.0169	0.0788
Group G2	0.0888	0.2087	0.0774	0.0155

Table S1*: The average distance of the curves within a cluster to the cluster average for each of the four groups for our optimal four-cluster realisation. This illustrates that for four groups within-cluster distances are less than the distances between clusters. For more clusters, within and between distances become similar.*

Region references	**H (+/-) *km***	**ρ-lc *kg/m^3***	**V_S-uc *km/s***	**V_S-lc *km/s***	**V_P-uc *km/s***	**V_P-lc *km/s***	**V_P/V_S used**	**V_P/V_S lit.**	**H_{uc}/H_{lc}**	**A_C int. *mW/m^2***
S1A[1,2,3]	38 (1)	3070.	3630	3960.	6500.	7190.	1.79-1.82	1.72	1.0	26.6
S1B[2,1]	42 (2)	3050	3630	3900	6400	6950	1.76-1.78	1.71	1.5	33.44
S2A[2]	44 (2)	3110	3620	4090	6600	7350	1.80-1.82	-	0.82	29.48
S2B[1,2,3]	46 (2)	3060	3650	3800	6600	6885	1.81	1.72	1.5	28.8
S2C[1,4]	36 (2)	3030	3590	3900	6300	6800	1.74-1.75	1.7-1.78	1.0	25.2
G1A[1,3]	40 (2)	2850.	3650.	3900.	6500.	6950.	1.78	1.68-1.82	1.0	28.0
G1B[1,4,5]	42 (2)	3030	3650.	3900	6400.	6950.	1.75-1.78	1.73-1.81	1.0	29.4
G2A[1,3(,2)]	40 (2)	3030	3550	3750.	6400	6900.	1.80-1.84	1.84	0.82	25.46
G2B[1,4,5]	41 (2)	3030	3670	4000	6500	7100	1.77-1.78	1.73-1.81	1.0	28.0

Table S2: *Crustal parameters, H – crustal thickness, ρ – density, uc – upper crust, lc – lower crust, V_P/V_S used – range of Vp/Vs in upper and lower crust, V_P/V_S lit. – range of V_P/V_S in literature, Huc/Hlc – ratio of upper over lower crustal thickness, A_c int – integrated heat production assigned to the crust in thermal modelling. Upper crustal density is 2850 kg/m^3 for all regions. References: 1: Petrescu et al. [2016], 2: Gilligan et al. [2016], 3: Berry and Fuchs [1973], 4: Winardhi and Mereu [1997], 5: Musacchio et al. [1997]*

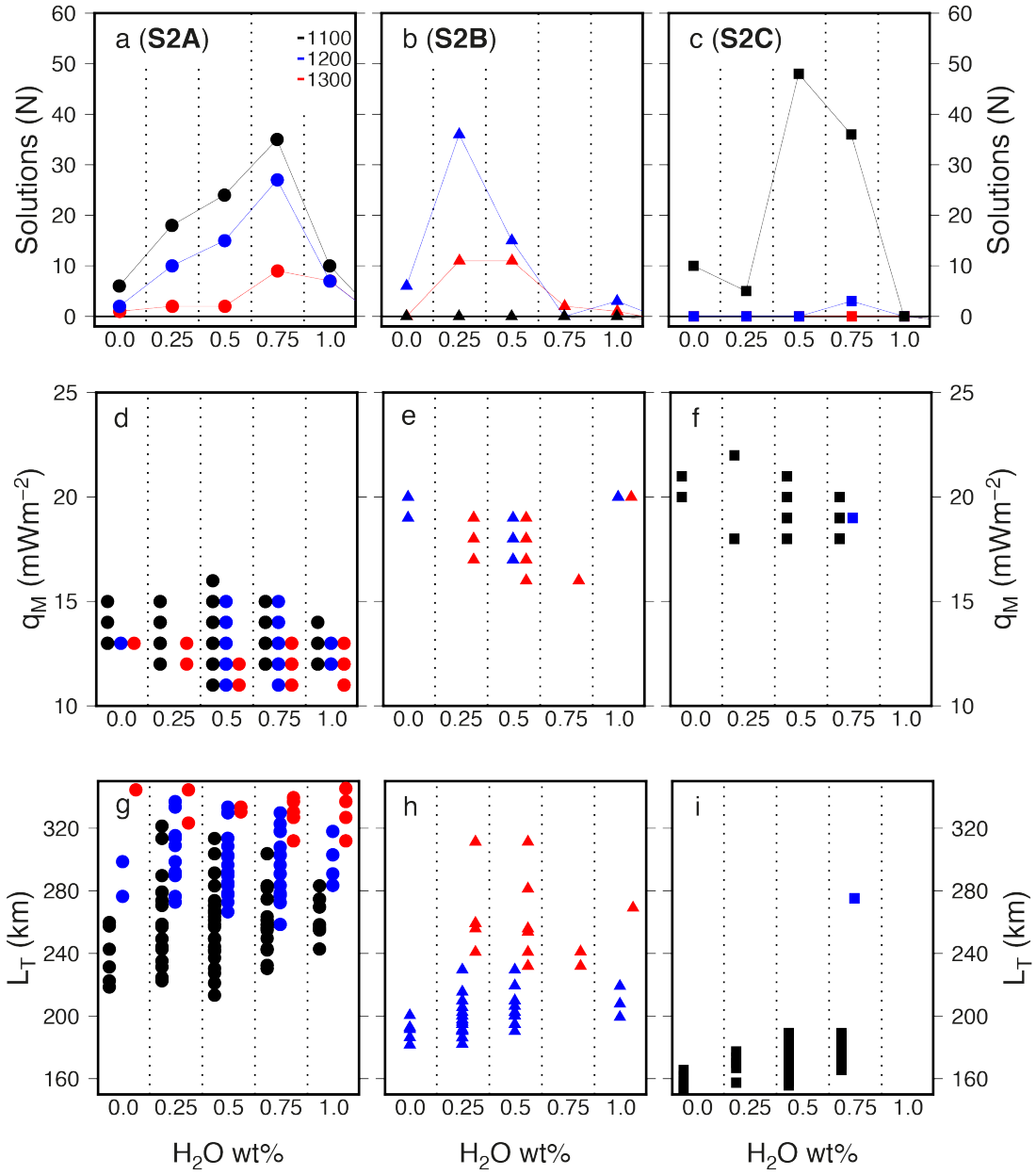

***Figure S2**: Summary of the results for the Superior subregions S2A, S2B and S2C for a **dunite background** composition, in terms of the number of solutions (a-c) and corresponding Moho heatflux, q_M, (d-f) and thermal lithospheric thicknesses, L_T, (g-h) as a function of the amount of water added. Compare to Figure 9, which shows results for a peridotite background composition. The different colours correspond to different potential temperatures for the asthenosphere (black - 1100°C, blue – 1200°C, red – 1300°C). Compared to a peridotitic background, slightly higher amounts of alteration are required.*

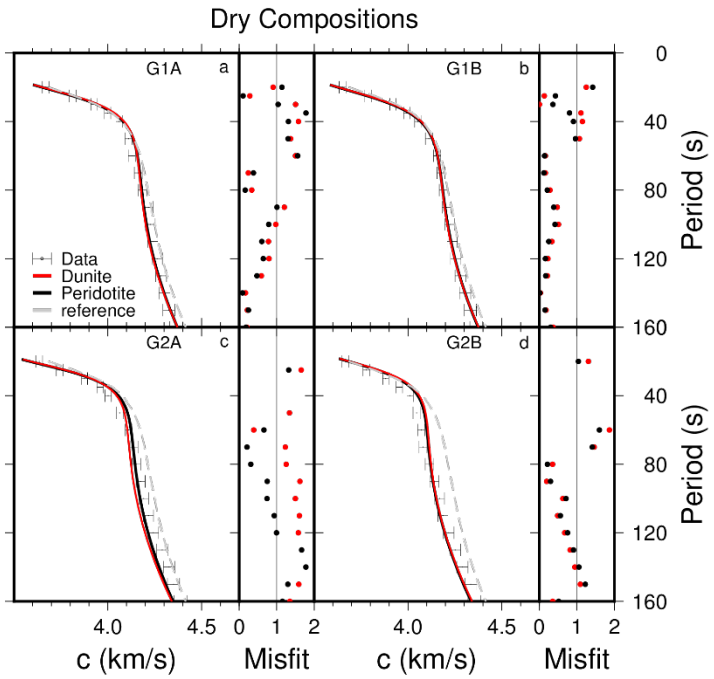

***Figure S3**: Lowest-misfit dispersion curves for a **dry dunite** and **dry peridotite** composition for all Grenville regions: G1A (**a**), and G1B (**b**), G2A (**c**), G2B (**d**) (all for 1300°C mantle adiabats) illustrating that dry compositions do not give any acceptable fits for the Grenville structures. Red lines and symbols represent dunite models and black lines and symbols peridotite models. The left panel in each case shows model dispersion curves compared with the cluster-averaged dispersion data (black error bars). The right panels show the point misfits of the different models, which at several periods exceed the acceptable limit of 1.*

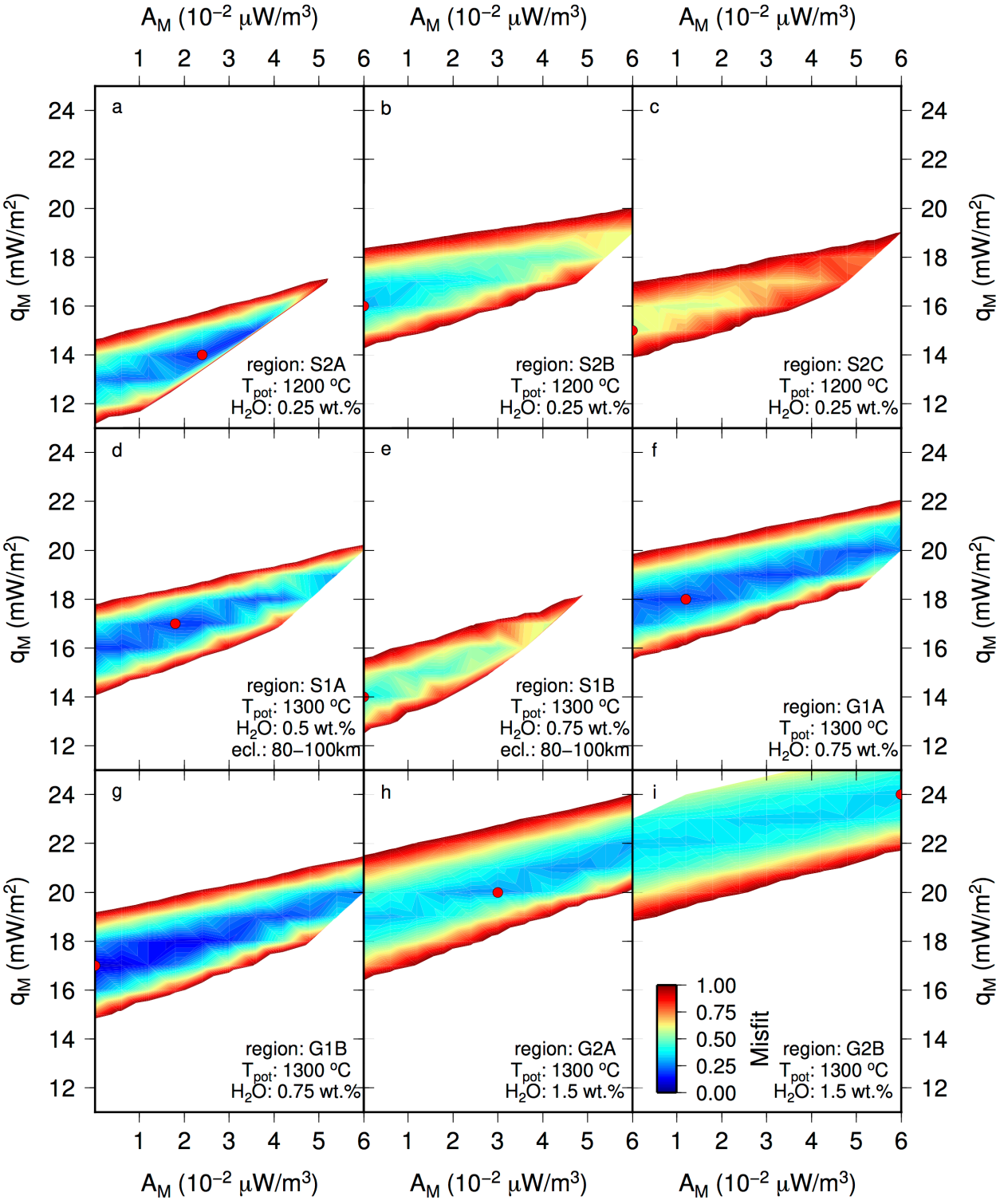

***Figure S4**: Thermal parameters of the solutions for all regions, for the water contents, eclogite layers and potential temperatures that yield the most solutions. The geotherm parameters displayed are the Moho heat flux q_M vs mantle heat production A_m, colour-coded according to their average misfits. The lowest average misfit solutions are indicated by the red dots, but all coloured solutions are acceptable.*

References:

Berry, M.J., Fuchs, K., 1973. Crustal structure of the superior and Grenville provinces of the Northeastern Canadian Shield. Bulletin of the Seismological Society of America 63, 1393-1432.

Gilligan, A., Bastow, I.D., Darbyshire, F.A., 2016. Seismological structure of the 1.8 Ga Trans-Hudson Orogen of North America. 17, 2421-2433, doi: 10.1002/2016GC006419.

Musacchio, G., Mooney, W.D., Luetgert, J.H., Christensen, N.I., 1997. Composition of the crust in the Grenville and Appalachian Provinces of North America inferred from V P /V S ratios. 102, 15225-15241, doi: 10.1029/96JB03737.

Petrescu, L., I. D. Bastow, F. A. Darbyshire, A. Gilligan, T. Bodin, W. Menke, and V. Levin (2016), Three billion years of crustal evolution in eastern Canada: Constraints from receiver functions, *121*(2), 788-811.

Rousseeuw, P. J. (1987), Silhouettes: A graphical aid to the interpretation and validation of cluster analysis, *Journal of Computational and Applied Mathematics*, *20*, 53-65.

Winardhi, S., Mereu, R.F., 1997. Crustal velocity structure of the Superior and Grenville provinces of the southeastern Canadian Shield. Canadian Journal of Earth Sciences 34, 1167-1184, doi: 10.1139/e17-094.