A comparison of global heat flow interpolation techniques

Buchanan C. Kerswell [1]**Matthew J. Kohn** [1]

[1]Department of Geosicences, Boise State University, Boise, ID 83725

Key Points:

-
-
-

Corresponding author: Buchanan C. Kerswell, buchanankerswell@u.boisestate.edu

Abstract

1 Introduction

Heat escaping the solid Earth's surface represents a cooling, active planet. Databases of surface heat flow (Hasterok & Chapman, 2008; Lucazeau, 2019; Pollack et al., 1993) provide a means of inferring lithospheric and mantle dynamics by relating the amount of heat escaping Earth's surface to diffusive heat loss (Fourier, 1827; Hasterok, 2013; Parsons & Sclater, 1977; Pollack & Chapman, 1977; Stein & Stein, 1992) and other heat-generating and heat-transferring processes Rudnick et al. (1998). Surface heat flow observations continue to motivate research, evident by more than 1,393 publications included in the most recent compilation (Jennings & Hasterok, 2021), although the rate of publications has declined since the mid 1980's.

Many research questions, such as calculating the global surface heat flux from continents and oceans, require interpolating discrete heat flow observations onto a continuous approximation of Earth's surface. Previous attempts used one or more geographic (e.g. bathymetry or elevation), geologic (e.g. proximity to active or ancient orogens), geochronologic (e.g. seafloor age), or geophysical (e.g. upper mantle shear wave velocities) proxies to associate and predict heat flow at unknown locations (Chapman & Pollack, 1975; Davies, 2013; Goutorbe et al., 2011; Lee & Uyeda, 1965; Lucazeau, 2019; Sclater & Francheteau, 1970; Shapiro & Ritzwoller, 2004). The success of interpolation methods are typically measured by the misfit between the predicted and observed heat flow. However, even statistically-successful heat flow interpolations are difficult to interpret and show unexpected anomalies (Lucazeau, 2019). The fidelity of interpolations are method-dependent and do not necessarily increase with decreasing misfit.

Predicting surface heat flow by association with physical proxies is arguably the most reasonable approach to this problem. Our understanding of geodynamics and near-surface heat flow perturbations implies that the variance in global surface heat flow is not stochastic, but rather determined by the set of physical conditions at point in space and processes operating locally around that point (e.g., Goutorbe et al., 2011). For example, younger oceanic plates should have higher surface heat flow than older plates (Stein & Stein, 1992), and surface heat flow may be modified from local ($\leq 60\ km$) hydrothermal circulation of seawater (Hasterok et al., 2011).

There must also exist some degree of stationarity, or spatial continuity, in the distribution of surface heat flow. For example, a pair of surface heat flow observations taken one meter apart will be strongly correlated. The correlation between pairs of observations will decrease with increasing distance between the pairs (Goovaerts & others, 1997). In theory, in the absence of near-surface perturbations, one may predict surface heat flow at unknown locations by considering nearby observations and their covariance (Kriging, Krige, 1951). The predictability of Kriging scales with the spatial coverage of observations. The method will be limited in areas with sparse data, but is arguably less artificial than methods of associating physical proxies to unknown heat flow locations used in previous studies (similarity methods, e.g., Goutorbe et al., 2011; Lucazeau, 2019; Shapiro & Ritzwoller, 2004).

The spatial continuity of surface heat flow is important because it allows one to asses the areal extent of potential thermal processes operating at depth. For example, patterns of surface heat flow sampled perpendicular to trenches of circum-Pacific subduction zones segments are relatively consistent from segment to segment (C. A. Currie & Hyndman, 2006; Furukawa, 1993; Wada & Wang, 2009). Consistency in heat flow patterns near arcs are interpreted to reflect common backarc lithospheric thermal structures and slab-mantle mechanical coupling depths (Furukawa, 1993; Kerswell et al., 2020; Wada & Wang, 2009).

This study uses ordinary Kriging to interpolate the New Global Heat Flow (NGHF) database of Lucazeau (2019). Our method is optimized using a Genetic Algorithm to minimize an objective function that considers both the misfit on the variogram models and interpolation results (e.g., Li et al., 2018). We then compare our interpolation results to those of Lucazeau (2019) and consider the implications of Kriging vs. similarity methods of interpolation. We restrict our comparison to areas near subduction zone segments defined by Syracuse & Abers (2006) for two reasons: 1) to provide maps and statistics useful to subduction zone research, and 2) to emphasize differences and idiosyncrasies in both interpolation approaches in a complex tectonic and thermal setting.

2 Methods

2.1 The NGHF Dataset

The NGHF dataset was downloaded from the supplementary material of Lucazeau (2019). It contains 69729 data points, their locations in latitude/longitude, and metadata—including a data quality rank (code6) from A to D (with code6 = Z = undetermined). The reader is referred to Lucazeau (2019) for details on compilation, references, and historical perspective on the NGHF and previous compilations. We use NGFH because it is the most recent dataset available, has been carefully compiled, and is open-access.

Like Lucazeau (2019), we exclude 4790 poor quality observations (code6 = D) from our analysis. We further remove 350 data points without heat flow observations and two without geographic information. Multiple observations at the same location are parsed to avoid singular covariance matrices during Krigings:

$$
\begin{aligned}
&f(X_i^q, Y_i^q) = \\
&X_i^q > Y_i^q \rightarrow z_i = x_i \\
&X_i^q < Y_i^q \rightarrow z_i = y_i \\
&X_i^q = Y_i^q \rightarrow z_i = RAND(x_i, y_i)
\end{aligned}
\tag{1}
$$

where X_i^q and Y_i^q represent the quality of each duplicate observation pair at location i, $RAND$ is a random function that selects either the observation x_i or y_i, and z_i stores the observation selected by $f(X_i^q, Y_i^q)$. The final dataset used for Kriging has $n = 55274$ observations after parsing $n = 32430$ duplicate observation.

2.2 Kriging

Kriging is a three-step process that involves first estimating an experimental variogram, $\hat{\gamma}(h)$, fitting the experimental variogram with one of many variogram models, $\gamma(h)$, and finally using the modelled variogram to predict random variables at unknown locations (Cressie, 2015; Krige, 1951). We use the general-purpose functions defined in the "R" package `gstat` (Gräler et al., 2016; Pebesma, 2004) to perform all three steps. We begin by estimating an experimental variogram as defined by Bárdossy (1997):

$$\hat{\gamma}(h) = \frac{1}{2N(h)} \sum_{N(h)} (Z(u_i) - Z(u_j))^2 \tag{2}$$

where $N(h)$ is the number of pairs of points, $Z(u_i)$ and $Z(u_j)$, separated by a lag distance, $h = |u_i - u_j|$. We evaluate $\hat{\gamma}(h)$ at fifteen lag distances by binning the irregular spaced data with a bin width, δ, equal to one-third of the maximum lag distance divided by the number of lags used to evaluate the variogram, $\delta = \max(N(h))/(3{\cdot}15)$. Then $N(h) \leftarrow N(h, \delta h) = \{i, j : |u_i - u_j| \in [h - \delta h, h + \delta h)\}$. In simple terms, Equation 2 represents the similarity, or dissimilarity, between pairs of observations in space. Equation 2 is derived from the theory of *regionalized variables* (Matheron, 1963, 2019), which formally defines a probabilistic framework for spatial interpolation of natural phenomena. It is important for the reader to understand the fundamental assumptions implicit in Equation 2 in order to understand the comparison of interpolation techniques discussed later. The basic assumptions used in our Kriging method are:

- $\hat{\gamma}(h)$ is directionally invariant (isotropic)
- $\hat{\gamma}(h)$ is evaluated in two-dimensions and neglects elevation, $Z(u) \in \mathbb{R}^2$
- The first and second moments of $Z(u)$ are spatially invariant in the domain D:

$$\begin{aligned} E[Z(u)] &= mean = constant, && \forall u \in D \\ E[(Z(u+h) - mean)(Z(u) - mean)] &= C(h), && \forall |u, u+h| \in D \end{aligned} \tag{3}$$

The last assumption (Equation 3) is called "second-order stationarity" and is commonly used in practice. It assumes the underlying probability distribution of the random variable, $Z(u)$, does not change in space and the covariance, $C(h)$, only depends on the distance, h, between two random variables. These assumptions are expected to be valid in cases where the underlying natural process is stochastic, spatially continuous, and has the property of additivity such that $\frac{1}{n}\sum_{i=1}^{n} Z(u_i)$ has the same meaning as $Z(u)$ (Bárdossy, 1997).

The following are two illustrative cases where Equation 3 is likely valid:

1. The thickness of a sedimentary unit with a homogeneous concentration of radioactive elements can be approximated by $q_s = q_b + \int A\, dz$, where q_b is a constant heat flux entering the bottom of the layer and A is the heat production within the layer with thickness z (Furlong & Chapman, 2013). If we have two samples, $Z(u_1) =$

31 mW/m^2 and $Z(u_2) = 30.5\ mW/m^2$, their corresponding thicknesses would be $Z'(u_1) = 1000\ m$ and $Z'(u_2) = 500\ m$ for $A = 0.001\ mW/m^3$ and $q_b = 30\ mW/m^2$. The variable, $Z(u)$, in this case is additive because the arithmetic mean of the samples is a good approximation of the average sedimentary layer thickness, $(Z(u_1) + Z(u_2))/2 = 750\ m$.

2. The age of young oceanic lithosphere can be approximated by $q_s(t) = kT_b(\pi\kappa t)^{-1/2}$, where $q_s(t)$ is the surface heat flow of a plate with age, t, T_b is the temperature at the base of the plate, k is thermal conductivity, and $\kappa = k/\rho C_p$ is thermal diffusivity (Stein & Stein, 1992). For $k = 3.138\ W/mK$, $\rho = 3330\ kg/m^3$, $C_p = 1171\ J/kgK$, $T_b = 1350°C$, two samples, $Z(u_1) = 180\ mW/m^2$ and $Z(u_2) = 190\ mW/m^2$, would correspond to plates with ages of $Z'(u_1) = 10\ Ma$, and $Z'(u_2) = 9\ Ma$, respectively. Since $Z(u_1)+Z(u_2)/2 = 185\ mW/m^2$ and $Z'(185\ mW/m^2) = 9.5\ Ma = Z'(u_1) + Z'(u_2)/2$, the variable $Z(u)$ in this case is also additive.

In contrast, Equation 3 is likely invalid in regions that transition among two or more tectonic regimes. For example, the expected heat flow $E[Z(u)] = mean$ will change when moving from a spreading center to a subduction zone. $E[Z(u)] = mean \neq constant$ over the region of interest. Proceeding with Equation 3 in this case has the effect of masking the geodynamic complexity. In other words, the spatial dependence is considered in the Kriging method in this case, but the geodynamic structure is *invisible*. We will see that this has important implications when comparing our Kriging method to Lucazeau (2019)'s interpolation method, which is exactly opposite of this formalism—it only considers the similarities among physical proxies and not spatial dependence.

The second step is to fit the experimental variogram with a variogram model, $\gamma(h)$. In this study we fit three popular variogram models to the experimental variogram. We use models with sills, which implies the spatial dependence between pairs of points has a finite range. The spherical, exponential, and Guassian variogram models are defined as (Chiles & Delfiner, 2009; Cressie, 2015):

$$
\begin{aligned}
sph \leftarrow \gamma(h) &= \begin{cases} n + s\left(\frac{3h}{2a} - \frac{1}{2}\left(\frac{h}{a}\right)^3\right), & \text{if } 0 \le h \le a \\ n + s, & \text{if } h > a \end{cases} \\
exp \leftarrow \gamma(h) &= n + s\left(1 - exp\left(\frac{-h}{a}\right)\right), \quad \text{if } h \ge 0 \\
Gau \leftarrow \gamma(h) &= n + s\left(1 - exp\left(\frac{-h^2}{a^2}\right)\right), \quad \text{if } h \ge 0
\end{aligned}
\tag{4}
$$

where n is the nugget, s is the sill, and a is the effective range. For spherical, exponential, and Gaussian models, the effective range is related to the range, r, by $a = r$, $a = r/3$, and $a = 1/r\sqrt{3}$, respectively (Gräler et al., 2016; Pebesma, 2004). The function `fit.variogram` in `gstat` allows one to try many variogram models and the best will be selected by the minimum misfit by weighted least square (WLS, Pebesma, 2004).

We use ordinary Kriging for our interpolation step, which predicts the value of a random function, $\hat{Z}(u)$ at unknown locations as a linear combination of all known locations in the domain, D (Bárdossy, 1997):

$$
\hat{Z}(u) = \sum_{i=1}^{n} \lambda_i Z(u_i), \quad \forall u \in D \tag{5}
$$

The conditions in Equation 3 set up a constrained minimization problem since one has:

$$
E[Z(u)] = mean, \quad \forall u \in D \tag{6}
$$

The linear estimator must obey

$$
E[\hat{Z}(u)] = \sum_{i=1}^{n} \lambda_i E[Z(u_i)] = mean \tag{7}
$$

so the weights must be:

$$
\sum_{i=1}^{n} \lambda_i = 1 \tag{8}
$$

This is the first constraint, also known as the unbiased condition, which states that the sum of the weights must equal one. However, there is an infinite set of real numbers

one could use for the weights, λ_i. Our goal is to find the set of weights in Equation 5 that minimizes the estimation variance. This can be solved with the covariance function, $C(h)$ from Equation 3:

$$\begin{gathered}\sigma^2(u) = Var[Z(u) - \hat{Z}(u)] = E\left[(Z(u) - \sum_{i=1}^{n} \lambda_i Z(u_i))^2\right] = \\ E\left[Z(u)^2 + \sum_{j=1}^{n}\sum_{i=1}^{n} \lambda_j \lambda_i Z(u_j) Z(u_i) - 2\sum_{i=1}^{n} \lambda_i Z(u_i) Z(u)\right] = \\ C(0) + \sum_{j=1}^{n}\sum_{i=1}^{n} \lambda_j \lambda_i C(u_i - u_j) - 2\sum_{i=1}^{n} \lambda_i C(u_i - u)\end{gathered} \tag{9}$$

Solving for the weights in Equation 5 with respect to the unbiased condition (Equation 8) and minimum estimate variance (Equation 9), yields the best linear unbiased estimator (BLUE, Bárdossy, 1997). In our case, this is done by the function `krige` in `gstat`.

2.3 Kriging Optimization

Achieving a useful Kriging results depends on one's choice of many Kriging parameters (Θ). In this study, we investigate a set of parameters, Θ:

$$\Theta = \{m, s, a, n, S\} \tag{10}$$

where m is the model type (sph, exp, or Gau), s is the sill, a is the effective range, n is the nugget, and S is the maximum distance for local Kriging. Only points within S from the prediction location are used for Kriging. Our goal is to find Θ such that our interpolation, $f(x_i; \Theta)$, gives the most useful outcome—defined by minimizing a cost function, $C(\Theta)$, that represents the error between the set of real observations, $Z(u_i)$ and predictions, $\hat{Z}(u)$.

We define a cost function that simultaneously considers the misfit between the experimental and modelled variogram and between the Kriging predictions and observed heat flow (after Li et al., 2018):

$$C(\Theta) = (1 - w) C_F(\Theta) + w C_I(\Theta) \tag{11}$$

where $C_F(\Theta)$ is the root mean square error (RMSE) of the modelled variogram fit calculated by WLS, and $C_I(\Theta)$ is the RMSE of the Kriging result calculated by cross-validation. The weight, w, is set to 0.5 in our study, which balances the effects of $C_F(\Theta)$ and $C_I(\Theta)$ on the cost function. The final expression to minimize becomes:

$$\min(C(\Theta)) = \frac{1-w}{\sigma_E}\sqrt{\frac{1}{N}\sum_{k=1}^{N} w(h_k)[\hat{\gamma}(h_k) - \gamma(h_k;\Theta)]^2} \quad + \\ \frac{w}{\sigma_S}\sqrt{\frac{1}{M}\sum_{i=1}^{M}[Z(u_i) - \hat{Z}(u_i;\Theta)]^2} \tag{12}$$

where N is the number of pairs of points used to calculate the experimental variogram, $\hat{\gamma}(h_k)$, σ_E is the standard deviation of the experimental variogram, $\hat{\gamma}(h)$, $w(h_k)$ is the weight in WLS and defines the importance of the kth lag in the error estimate. We use $w(h_k) = N_k/h_k^2$. $Z(u_i)$ and $\hat{Z}(u_i;\Theta)$ are the measured and predicted values, respectively, σ_s is the standard deviation of the predicted values, $\hat{Z}(u_i)$, and M is the number of measurements in $Z(u_i)$. For $C_I(\Theta)$ we use ten-fold cross-validation, which splits the dataset, $|Z(u_i),\ \forall u_i \in D|$ into ten equal intervals and tests one interval against the remaining nine. This process is then repeated over all intervals so that the whole dataset has been cross-validated.

Minimization of $C(\Theta)$ is achieved by a genetic algorithm that simulates biologic natural selection by differential success (Goldberg, 1989). Our procedure is as follows:

1. Initiate fifty *chromosomes*, ξ, with random starting parameters defined within the search domain (Table 1)
2. Evaluate the fitness of each individual chromosome as $-C(\Theta)$ for the entire population
3. Allow the population to exchange genetic information by sequentially performing genetic operations:
   a. Selection: the top 5% fittest chromosomes survive each generation
   b. Crossover: pairs of chromosomes have an 80% chance of exchanging genetic information
   c. Mutation: there is a 10% chance for random genetic mutations
4. Evaluate the fitness of the new population

5. If the termination criterion is met, do step (6), otherwise continue to evolve by repeating steps (3) and (4)
6. Decode the best chromosome and build the optimal variogram

We use the general-purpose functions in the "R" package `GA` (Scrucca, 2013, 2016) to perform each step in the above procedure.

Table 1: Parameters and ranges used in the optimization algorithm

Parameter	Search Domain	Units
Model (m)	[Spherical, Exponential, Guassian]	NA
Sill (s)	$[1, 5 \times 10^3]$	$\left(mW/m^2\right)^2$
Effective Range (a)	$[1, 1 \times 10^6]$	meters
Nugget (n)	$[1, 1 \times 10^3]$	meters
Local Search (S)	$[1, 1 \times 10^6]$	meters

2.4 Map Projection and Interpolation Grid

We interpolate onto the same 0.5°C x 0.5°C grid as Lucazeau (2019) so a direct difference could be calculated between our interpolation methods and Lucazeau (2019)'s. The NGHF and grid with predicted heat flow from Lucazeau (2019) were transformed into a Pacific-centered Robinson coordinate reference system (CRS) defined using the `proj` string (PROJ contributors, 2021):

```
+proj=robin +lon_0=-155 +lon_wrap=-155 +x_0=0 +y_0=0
+ellps=WGS84 +datum=WGS84 +units=m +no_defs
```

All geographic operations, including Kriging, are performed in the above CRS using the general-purpose functions in the "R" package `sf` (Pebesma, 2018). We provide the complete NGHF dataset (Lucazeau, 2019), filtered and parsed NGHF dataset, heat flow interpolations (from Lucazeau, 2019, and this study), and our code as supplementary information to support FAIR data policy (Wilkinson et al., 2016). These items can also be retrieved from the official repository at `https://doi.org/10.17605/OSF.IO/CA6ZU`.

222 **3 Results**

223 **4 Discussion**

224 **5 Conclusions**

225 **Acknowledgments**

References

Bárdossy, A. (1997). Introduction to geostatistics. *Institute of Hydraulic Engineering, University of Stuttgart*.

Chapman, D. S., & Pollack, H. N. (1975). Global heat flow: A new look. *Earth and Planetary Science Letters*, *28*(1), 23–32.

Chiles, J.-P., & Delfiner, P. (2009). *Geostatistics: Modeling spatial uncertainty* (Vol. 497). John Wiley & Sons.

Cressie, N. (2015). *Statistics for spatial data*. John Wiley & Sons.

Currie, C., Wang, K., Hyndman, R. D., & He, J. (2004). The thermal effects of steady-state slab-driven mantle flow above a subducting plate: The cascadia subduction zone and backarc. *Earth and Planetary Science Letters*, *223*(1-2), 35–48.

Currie, C. A., & Hyndman, R. D. (2006). The thermal structure of subduction zone back arcs. *Journal of Geophysical Research: Solid Earth*, *111*(B8).

Davies, J. H. (2013). Global map of solid earth surface heat flow. *Geochemistry, Geophysics, Geosystems*, *14*(10), 4608–4622.

Fourier, J. (1827). Mémoire sur les températures du globe terrestre et des espaces planétaires. *Mémoires de l'Académie Royale Des Sciences de l'Institut de France*, *7*, 570–604.

Furlong, K. P., & Chapman, D. S. (2013). Heat flow, heat generation, and the thermal state of the lithosphere. *Annual Review of Earth and Planetary Sciences*, *41*, 385–410.

Furukawa, Y. (1993). Depth of the decoupling plate interface and thermal structure under arcs. *Journal of Geophysical Research: Solid Earth*, *98*(B11), 20005–20013.

Gao, X., & Wang, K. (2014). Strength of stick-slip and creeping subduction megathrusts from heat flow observations. *Science*, *345*(6200), 1038–1041.

Goldberg, D. E. (1989). Genetic algorithms in search. *Optimization, and MachineLearning*.

Goovaerts, P., & others. (1997). *Geostatistics for natural resources evaluation*. Oxford University Press on Demand.

Goutorbe, B., Poort, J., Lucazeau, F., & Raillard, S. (2011). Global heat flow trends resolved from multiple geological and geophysical proxies. *Geophysical Journal International*, *187*(3), 1405–1419.

Gräler, B., Pebesma, E., & Heuvelink, G. (2016). Spatio-temporal interpolation using gstat. *The R Journal*, *8*, 204–218. Retrieved from `https://journal.r-project.org/archive/2016/RJ-2016-014/index.html`

Hasterok, D. (2013). A heat flow based cooling model for tectonic plates. *Earth and Planetary Science Letters*, *361*, 34–43.

Hasterok, D., & Chapman, D. (2008). Global heat flow: A new database and a new approach. In *AGU fall meeting abstracts* (Vol. 2008, pp. T21C–1985).

Hasterok, D., Chapman, D., & Davis, E. (2011). Oceanic heat flow: Implications for global heat loss. *Earth and Planetary Science Letters*, *311*(3-4), 386–395.

Jennings, S., & Hasterok, D. (2021). HeatFlow.org. *Heatflow.org*. Retrieved from `http://heatflow.org/`

Kerswell, B. C., Kohn, M. J., & Gerya, T. V. (2020). Backarc lithospheric thickness and serpentine stability control slab-mantle coupling depths in subduction zones. *Earth and Space Science Open Archive*, 34. `https://doi.org/10.1002/essoar.10503710.1`

Krige, D. G. (1951). A statistical approach to some basic mine valuation problems on the witwatersrand. *Journal of the Southern African Institute of Mining and Metallurgy*, *52*(6), 119–139.

Lee, W. H., & Uyeda, S. (1965). Review of heat flow data. *Terrestrial Heat Flow*, *8*, 87–190.

Li, Z., Zhang, X., Clarke, K. C., Liu, G., & Zhu, R. (2018). An automatic variogram modeling method with high reliability fitness and estimates. *Computers & Geosciences*, *120*, 48–59.

Lucazeau, F. (2019). Analysis and mapping of an updated terrestrial heat flow data set. *Geochemistry, Geophysics, Geosystems*, *20*(8), 4001–4024.

Matheron, G. (1963). Principles of geostatistics. *Economic Geology*, *58*(8), 1246–1266.

Matheron, G. (2019). *Matheron's theory of regionalized variables*. International Association for.

Parsons, B., & Sclater, J. G. (1977). An analysis of the variation of ocean floor bathymetry and heat flow with age. *Journal of Geophysical Research*, *82*(5), 803–827.

Pebesma, E. (2004). Multivariable geostatistics in S: The gstat package. *Computers & Geosciences*, *30*, 683–691.

Pebesma, E. (2018). Simple Features for R: Standardized Support for Spatial Vector Data. *The R Journal*, *10*(1), 439–446. https://doi.org/10.32614/RJ-2018-009

Pollack, H. N., & Chapman, D. S. (1977). On the regional variation of heat flow, geotherms, and lithospheric thickness. *Tectonophysics*, *38*(3-4), 279–296.

Pollack, H. N., Hurter, S. J., & Johnson, J. R. (1993). Heat flow from the earth's interior: Analysis of the global data set. *Reviews of Geophysics*, *31*(3), 267–280.

PROJ contributors. (2021). *PROJ coordinate transformation software library*. Open Source Geospatial Foundation. Retrieved from https://proj.org/

Rudnick, R. L., McDonough, W. F., & O'Connell, R. J. (1998). Thermal structure, thickness and composition of continental lithosphere. *Chemical Geology*, *145*(3-4), 395–411.

Sclater, J. G., & Francheteau, J. (1970). The implications of terrestrial heat flow observations on current tectonic and geochemical models of the crust and upper mantle of the earth. *Geophysical Journal International*, *20*(5), 509–542.

Scrucca, L. (2013). GA: A package for genetic algorithms in r. *Journal of Statistical Software*, *53*(4), 1–37.

Scrucca, L. (2016). On some extensions to GA package: Hybrid optimisation, parallelisation and islands evolution. *arXiv Preprint arXiv:1605.01931*.

Shapiro, N. M., & Ritzwoller, M. H. (2004). Inferring surface heat flux distributions guided by a global seismic model: Particular application to antarctica. *Earth and Planetary Science Letters*, *223*(1-2), 213–224.

Stein, C. A., & Stein, S. (1992). A model for the global variation in oceanic depth and heat flow with lithospheric age. *Nature*, *359*(6391), 123–129.

Stein, C. A., & Stein, S. (1994). Constraints on hydrothermal heat flux through the oceanic lithosphere from global heat flow. *Journal of Geophysical Research: Solid Earth*, *99*(B2), 3081–3095.

Syracuse, E. M., & Abers, G. A. (2006). Global compilation of variations in slab depth beneath arc volcanoes and implications. *Geochemistry, Geophysics, Geosystems*, *7*(5).

Wada, I., & Wang, K. (2009). Common depth of slab-mantle decoupling: Reconciling diversity and uniformity of subduction zones. *Geochemistry, Geophysics, Geosystems*, *10*(10).

Wilkinson, M. D., Dumontier, M., Aalbersberg, Ij. J., Appleton, G., Axton, M., Baak, A., et al. (2016). The FAIR guiding principles for scientific data management and stewardship. *Scientific Data*, *3*(1), 1–9.