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

```
## Loading libraries:
## magrittr
## ggplot2
## tidyr
## readr
## purrr
## gstat
## ggsflabel
## sf
## ggrepel
## patchwork
## cowplot
## dplyr
## Loading functions
```

1 Introduction

Heat escaping the solid Earth's surface indicates a dynamically cooling planet. Surface heat flow databases (Hasterok & Chapman, 2008; Lucazeau, 2019; Pollack et al., 1993) provide a way to understand geodynamics by relating the amount of heat escaping Earth's surface to heat-transferring and heat-generating subsurface processes such as diffusion, hydrothermal circulation, radioactive decay, fault motion, subduction dynamics, and mantle convection (Currie et al., 2004; Currie & Hyndman, 2006; Fourier, 1827; Furlong & Chapman, 2013; Furukawa, 1993; Gao & Wang, 2014; Hasterok, 2013; Kerswell et al., 2020; Parsons & Sclater, 1977; Pollack & Chapman, 1977; Rudnick et al., 1998; Stein & Stein, 1992, 1994; Wada & Wang, 2009). Surface heat flow observations continue to motivate research, evident by more than 1,393 publications compiled in the most recent heat flow dataset, although the rate of publications using surface heat flow has declined since the mid 1980's (Jennings & Hasterok, 2021).

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 at interpolation use one or more geographic, geologic, geochronologic, or geophysical proxies to predict heat flow at unknown locations by association with similar observation sites (e.g., bathymetry or elevation, proximity to active or ancient orogens, seafloor age, upper mantle shear wave velocities, Chapman & Pollack, 1975; Davies, 2013; Goutorbe et al., 2011; Lee & Uyeda, 1965; Lucazeau, 2019; Sclater & Francheteau, 1970; Shapiro & Ritzwoller, 2004). These methods are called *similarity methods* (Figure 1). The success of such interpolations are typically evaluated statistically 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 and usefulness of interpolations depend on the question being asked and the choice of methodology.

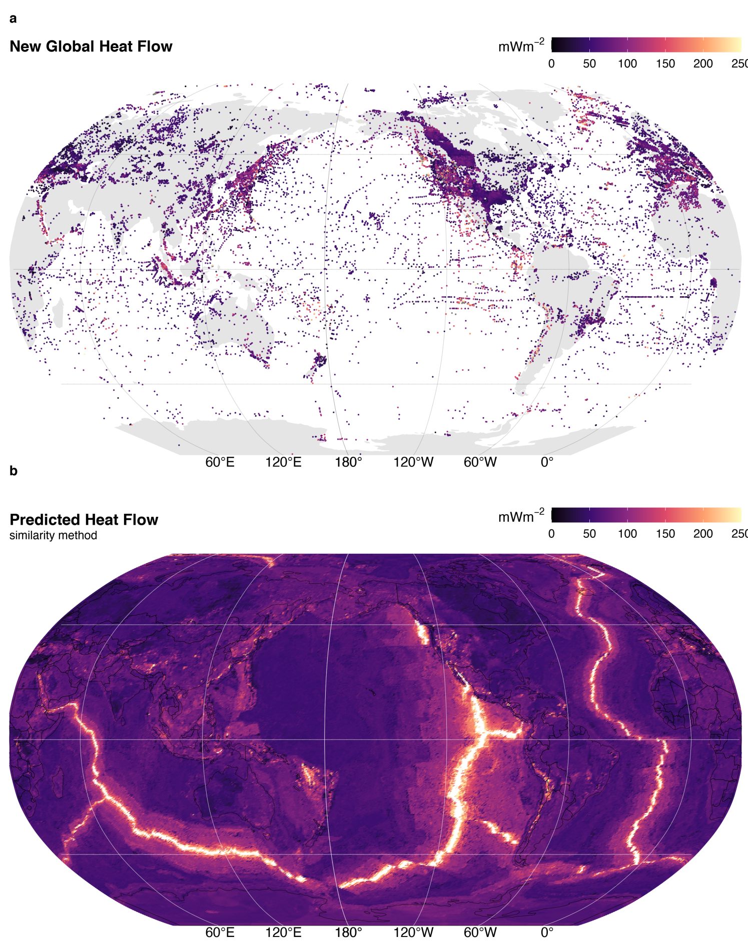

Figure 1: The NGHF dataset. (a) The complete dataset (n = 69729), and (b) interpolation by similarity method from Lucazeau (2019).

Predicting surface heat flow by association with physical proxies is arguably the most reasonable approach to interpolation for global investigations. Our understanding of geodynamics and near-surface heat flow perturbations implies that the variance in surface heat flow is not uniformly stochastic, but rather, in large part, determined by the physical conditions and processes operating locally (e.g., Goutorbe et al., 2011). For example, younger oceanic plates should have higher surface heat flow than older plates (Stein & Stein, 1992), subducting oceanic plates will lower surface heat flow near trenches (Furukawa, 1993), and hydrothermal circulation of seawater can modify heat flow in oceanic crust (Hasterok et al., 2011). Interpolation by association with physical proxies makes reasoned predictions of heat flow based on many independently-tested geodynamic models. However, similarity methods are strongly biased towards such models and risk making determinations where, in fact, surprising results and idiosyncrasies may be found.

In contrast, there exists some degree of stationarity, spatial dependence, or continuity, in the distribution of surface heat flow. A pair of surface heat flow observations taken one meter apart will be strongly correlated. The correlation between pairs of observations will likely decrease with increasing distance between the pairs (Goovaerts, 1997). The spatial (dis)continuity of surface heat flow represents the areal extent of geodynamic processes and their interactions. For example, consistent patterns of heat flow near volcanic arcs are interpreted to reflect common backarc lithospheric thermal structures and slab-mantle mechanical coupling depths in subduction zones (Furukawa, 1993; Kerswell et al., 2020; Wada & Wang, 2009).

In theory, one may predict surface heat flow at unknown locations by considering many nearby observations (i.e. Kriging, Krige, 1951). However, Kriging is disadvantageous for global interpolations of surface heat flow because it assumes that the underlying distribution of heat flow is stationary (constant in space), which effectively ignores geodynamic complexity. One can overcome this problem by relaxing assumptions of stationarity, or applying Markov-Bayes techniques to include proxies as priors (Bárdossy, 1997). Instead, we leverage the properties of stationarity as a tool for comparison with *similarity* methods of interpolation (Goutorbe et al., 2011; Lucazeau, 2019). So the questions are: 1) What are the differences between Kriging and similarity methods? 2) What are the implications of the differences according to the implicit assumptions in both methods?

We attempt to answer these questions by using ordinary Kriging to interpolate the New Global Heat Flow (NGHF) dataset of Lucazeau (2019). Our method is optimized using a genetic algorithm to minimize an objective function which considers both the misfit on the variogram models and interpolation results (after 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 (Code 6) from A to D (with Code 6 = 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 (Code 6 = 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 Kriging:

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

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 $Z(u)$ have the following conditions over 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 two 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 and exponential variogram models used in this study are defined as (Chiles & Delfiner, 2009; Cressie, 2015):

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

where n is the nugget, s is the sill, and a is the effective range. The effective range, a, is related to the range, r, by $a = r$ and $a = r/3$ for spherical and exponential models, respectively (Gräler et al., 2016; Pebesma, 2004). We use the function `fit.variogram` in `gstat` to try both variogram models. The best model is 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:

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

where c is the lag cutoff proportion, w is the lag window, m is the model type (sph or exp), 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. The lag cutoff is the maximum separation distance between pairs of points used in the experimental variogram (i.e. the x-axis maximum limit) calculated as a fraction of the overall maximum separation distance for all observations, $Z(u)$, in the domain, D. The lag window, w, shifts the lags where the variogram is evaluated by removing the first n lags and adding n lags to the right side of the variogram. This is necessary to avoid negative ranges, a, when fitting experimental variograms with anomalously high variances at small lag distances.

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):

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

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:

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

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
Lag Cutoff (c)	[1/3, 1/15]	NA
Lag Window (w)	[1, 5]	NA
Model (m)	[Spherical, Exponential]	NA
Sill (s)	$[1, 1000\sqrt{2}]$	mWm^{-2}
Effective Range (a)	[1, 1000]	km
Nugget (n)	$[1, 1000\sqrt{2}]$	mWm^{-2}
Local Search (S)	[1, 1000]	km

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 and taking the difference with Lucazeau (2019)'s heat flow predictions, are performed in the above CRS using the general-purpose functions in the "R" package `sf` (Pebesma, 2018). We define the Kriging domain near individual arc segments in two steps: 1) 1000 km buffers are drawn around the arc segments as defined by Syracuse & Abers (2006). 2) The bounding box of the 1000 km buffer is expanded by 10% on all sides (Figure 2). We use Lucazeau (2019)'s grid for Kriging predictions so differences can be taken point-by-point at the exact same locations.

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 materials can also be retrieved from the official repository at `https://doi.org/10.17605/OSF.IO/CA6ZU`.

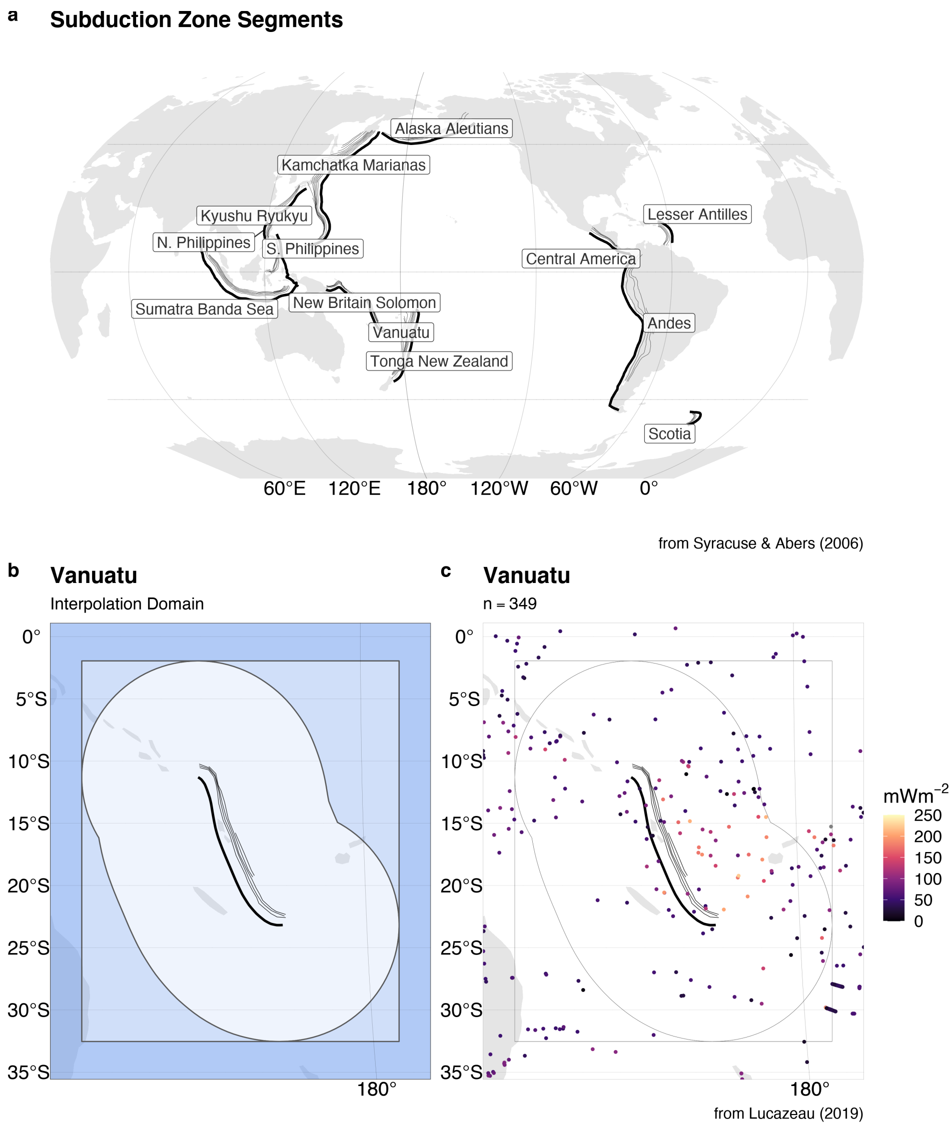

Figure 2: Subduction zone segments and interpolation domain. (a) Heat flow is interpolated around thirteen subduction zone segments by (b) drawing a 1000km buffer (lightest blue) around each segment and expanding the buffer's bounding box (medium blue) by 10% on all sides (darkest blue). (c) The NGHF dataset is cropped within the largest rectangle. Data from Syracuse & Abers (2006) and Lucazeau (2019).

3 Results

3.1 Heat Flow Near Subduction Zone Segments

Summary statistics for surface heat flow observations by subduction zone segment are given in Table 2 and Figure 3. Surface heat flow is median-centered around 45-70 mWm^{-2} and narrowly distributed (excluding outliers) with inter-quartile ranges (IQR) from 12 to 50 mWm^{-2} for most subduction zone segments. Alaska Aleutians is the exception with a higher median of 184 mWm^{-2} and broader range (IQR = 250 mW^{-2}). The whole distributions (including outliers) for all segments are strongly right-skewed with maximum heat flow values of several thousand of mWm^{-2} or more. Heat flow values above 250 mWm^{-2} are considered geothermal areas by Lucazeau (2019), which we adopt as a relevant empirical limit for anomalously high heat flow.

Table 2: Heat flow (mWm^{-2}) observations

Segment	n	Min	Max	Median	IQR	Mean	Sigma
Alaska Aleutians	2792	4	7765	184	206	233	344
Andes	5226	7	911	45	50	87	88
Central America	5038	8	911	45	48	88	91
Kamchatka Marianas	3956	1	31000	71	48	121	658
Kyushu Ryukyu	3246	1	31000	72	44	130	726
Lesser Antilles	3534	13	1150	41	12	52	50
N. Philippines	1001	3	31000	72	35	206	1276
New Britain Solomon	181	3	174	65	42	69	30
S. Philippines	1442	1	31000	71	43	167	1063
Scotia	72	13	145	76	23	74	28
Sumatra Banda Sea	3037	1	5600	65	47	77	120
Tonga New Zealand	507	1	416	54	40	64	45
Vanuatu	349	1	283	54	40	64	44

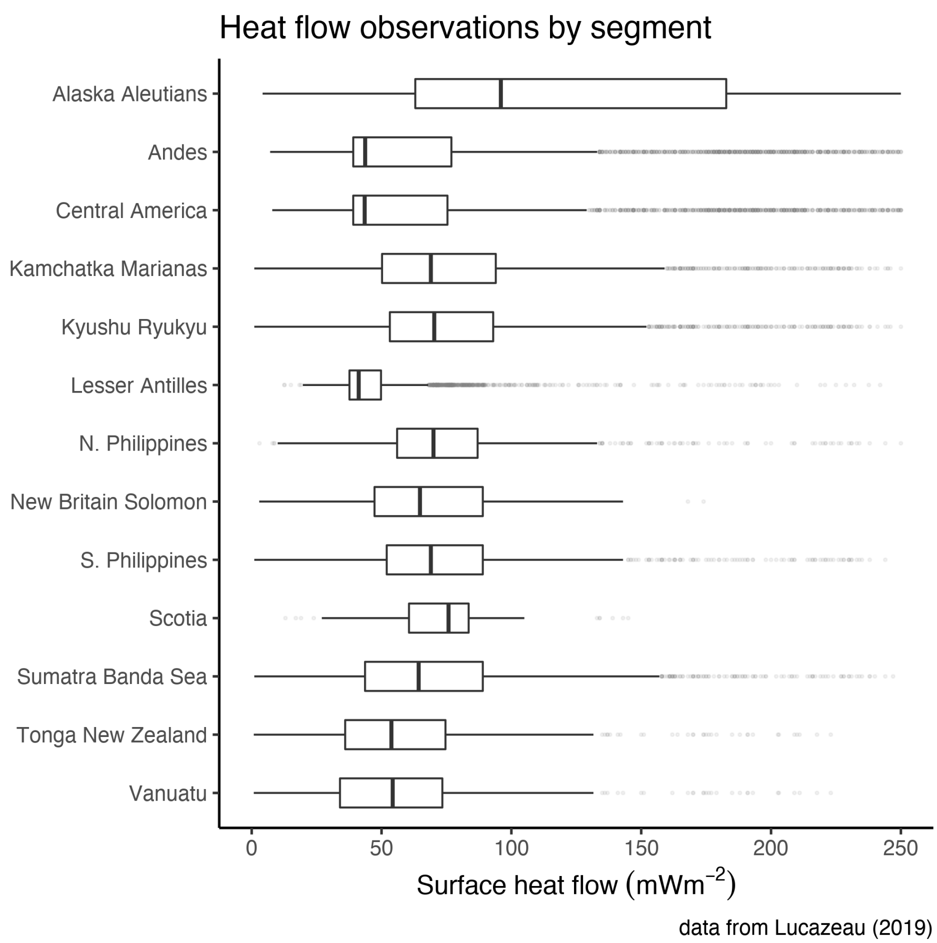

Figure 3: Distribution of heat flow observations. Heat flow near most segments is centered around 50 mWm^{-2} and highly skewed right (shadowy dot outliers). The skeweness likely represents sampling near geothermal systems, volcanic arcs, or spreading centers. Data from Lucazeau (2019).

3.2 Variogram Models

The optimal variogram models and associated errors $C_F(\Theta)$ and $C_I(\Theta)$ are given in Table 3. Almost twice as many experimental variograms are fit with spherical models (8) compared to exponential models (5). Variogram model sills vary substantially among the subduction zone segments between 9 and 1538 mWm^{-2}. Variogram model ranges also vary substantially among segments from 4 to 1676 km.

No apparent correlation exists between variogram model range and subduction zone segment length, number of heat flow observations, nor domain area (Figure 4). Most subduction zone segments show spatial dependence between 9 and $\leq$ 500 km, except for Kyushu Ryukyu and Vanuatu, irrespective of the amount of observations or segment size.

Table 3: Optimal varigram models

Segment	Model	Sill [mWm^{-2}]	Range [km]
Alaska Aleutians	Sph	284	219
Andes	Sph	87	344
Central America	Sph	86	342
Kamchatka Marianas	Exp	692	274
Kyushu Ryukyu	Sph	911	1676
Lesser Antilles	Exp	37	98
N. Philippines	Sph	1358	243
New Britain Solomon	Exp	28	76
S. Philippines	Sph	1163	421
Scotia	Exp	9	4
Sumatra Banda Sea	Exp	113	32
Tonga New Zealand	Sph	47	131
Vanuatu	Sph	48	574

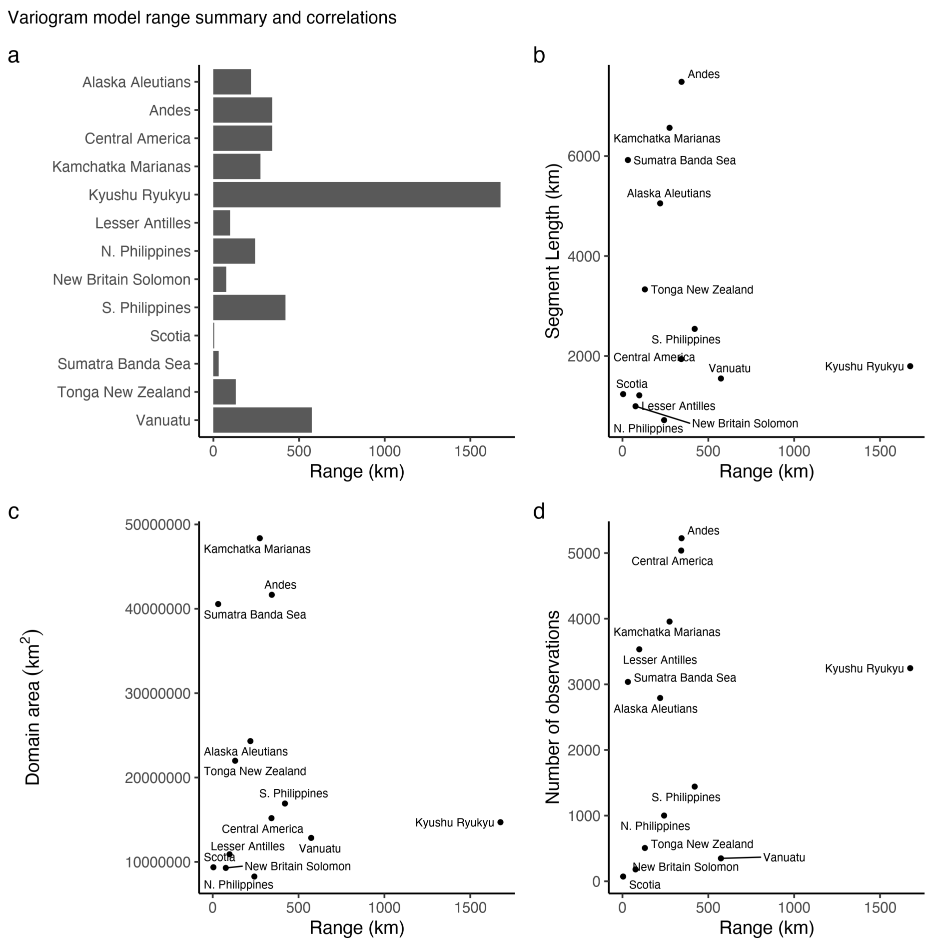

Figure 4: Summary of variogram model ranges and correlations with other features. (a) Variogram model ranges are variable, but generally below 500 km. Variogram model ranges show no correlation with segment length (b), number of heat flow observations (c), nor domain area (d). The spatial dependenc of heat flow is apparently independent of these parameters.

3.3 Interpolation Comparison

Summary statistics for the interpolation differences are given in Table 4 and Figure 5.

Table 4: Predicted heat flow (mWm^{-2}) differences

Segment	Min	Max	Median	IQR	Mean	Sigma
Alaska Aleutians	-550	2690	-2	20	3	71
Andes	-188	2336	2	35	13	72
Central America	-189	2731	8	62	44	146
Kamchatka Marianas	-2777	297	4	15	4	45
Kyushu Ryukyu	-2781	168	5	24	2	85
Lesser Antilles	-103	572	0	16	6	31
N. Philippines	-2800	278	0	26	-5	94
New Britain Solomon	-81	240	1	22	5	20
S. Philippines	-274	375	5	26	7	29
Scotia	-62	1001	-4	24	5	33
Sumatra Banda Sea	-243	385	-2	20	1	19
Tonga New Zealand	-113	1695	-9	23	-2	30
Vanuatu	-166	1657	7	27	8	40

4 Discussion

Variogram model ranges characterize the maximum separation distance at which pairs of points are spatially dependent (Matheron, 1963). Therefore, the range is the most important parameter to evaluate when considering spatial continuity of heat flow. # Conclusions

Acknowledgments

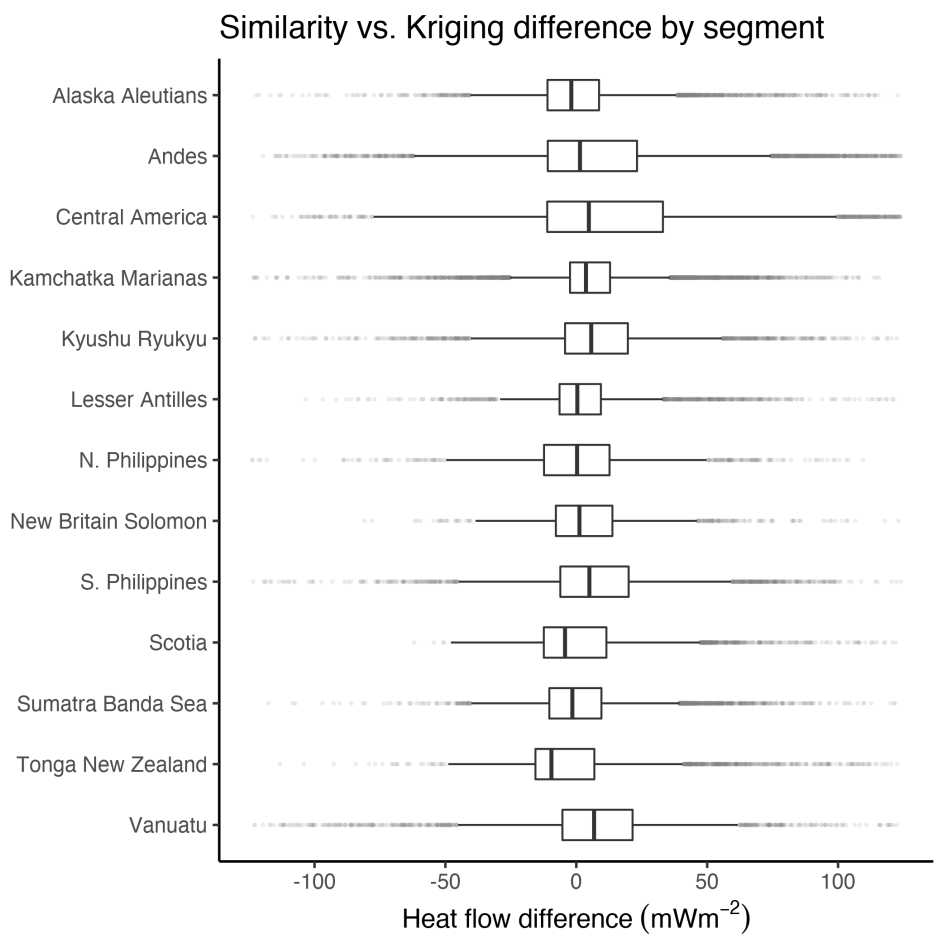

Figure 5: Point-by-point differences of predicted heat flow between similarity and Kriging interpolations. The differences for most subduction zone segments are median-centered at or near-zero with IQRs from 16 to 62. Outliers (shadowy dots) extend to extreme positive and negative differences.

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

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.

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. (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.