

¹ Antarctic bed topography estimation using a Stochastic Meshless ² Uncertainty Gridding (SMUG) method

³ Lenneke M. Jong^a, Jason L. Roberts^a, Felicity S. McCormack^b and Ana Fabela Hinojosa^b

⁴ *^aAustralian Antarctic Division, Department of Climate Change, Energy, the Environment and Water, Australia*

⁵ *^bSecuring Antarctica's Environmental Future, School of Earth, Atmosphere and Environment, Monash University, Clayton, Kulin Nations,
6 Victoria, Australia*

⁸ ARTICLE INFO

¹¹ **Keywords:**
Interpolation
spatial data
Antarctica
bed topography

ABSTRACT

An ongoing challenge in spatial sciences is the generation of gridded datasets for sparsely sampled data, with robust estimates of uncertainty. This is particularly the case for datasets of the bed topography beneath the Antarctic Ice Sheet — an area of approximately 12.1 million km² — which suffers from sampling biases due to the difficulty of directly observing this vast and remote region. Here, we present a new method for interpolating sparsely and unevenly sampled data: the Stochastic Meshless Uncertainty Gridding (SMUG) method. SMUG uses a point-by-point meshing scheme and Taylor series approximations to generate a large ensemble of estimates of the data field and its corresponding uncertainty at the interpolation point. We apply the SMUG method to estimate the bed topography of the Aurora Subglacial Basin, East Antarctica, finding high skill in the SMUG estimates when compared with existing gridded products. The robustness of the method was assessed by progressively excluding an increasing volume of observational data demonstrating that the SMUG method offers significant benefit for applications involving sparse datasets where robust uncertainty quantification is critical for informed decision-making and risk assessment.

25

²⁶ CRediT authorship contribution statement

²⁷ **Lenneke M. Jong:** Conceptualisation, Investigation, Writing - Original Draft . **Jason L. Roberts:** Conceptuali-
²⁸ sation, Software, Methodology, Investigation, Writing - Original Draft. **Felicity S. McCormack:** Conceptualisation,
²⁹ Investigation, Writing - Original Draft. **Ana Fabela Hinojosa:** Investigation, Writing - Original Draft.

³⁰ 1. Introduction

³¹ A key challenge in spatial sciences is in constructing datasets which are representative of continuous fields, using
³² discrete and often sparse observations and interpolation. The choice of an interpolation method which is able to
³³ capture the underlying variability of the field is not necessarily straightforward. This is a particular challenge when
³⁴ considering the bed topography of the Antarctic Ice Sheet, where the logistical challenge of observations means it
³⁵ remains sparsely sampled. At present, for a 500m by 500m horizontal grid, less than 7% of Antarctica's topography
³⁶ across the 12.10 million km² of grounded ice area contains observations (Pritchard et al., 2025), and there are large data
³⁷ gaps in dynamic regions along the grounding line. This has implications for the propagation of uncertainties in model
³⁸ projections of Antarctica's evolution under climate change, including the timing and rate of any potential grounding
³⁹ line retreat (McCormack et al., Submitted).

ORCID(s): 0000-0001-6707-570X (L.M. Jong); 0000-0002-3477-4069 (J.L. Roberts); 0000-0002-2324-2120 (F.S. McCormack);
0009-0000-7238-5761 (A.F. Hinojosa)

40 Several established interpolation methods have been used to generate Antarctic topography data sets. Here, we
 41 focus on methods that produce elevation-preserving bed topographies, i.e., where the measurements of the bed topog-
 42 raphy are retained in the gridded dataset produced by the interpolation, rather than methods that preserve other char-
 43 acteristics, such as texture or roughness (e.g. Graham et al., 2017; MacKie and Schroeder, 2020; MacKie et al., 2021).
 44 One of the earliest gridded bed topography datasets for Antarctica – Bedmap1 (Lythe and Vaughan, 2001) – employed
 45 an Inverse Distance Weighting (IDW) algorithm. This method estimates values at unsampled points as weighted aver-
 46 ages of nearby measurements, with weights inversely proportional to a power of distance (e.g., $weight = 1/distance^p$,
 47 where p is the power parameter, (Mitas and Mitasova, 1999)). In Bedmap1 the IDW algorithm uses the nearest two
 48 observations in each octant (a total of 16 data points) to estimate the value at each grid node. The octagonal approach
 49 was specifically designed to address the highly anisotropic distribution of the ice thickness survey data. After testing
 50 different weightings, Lythe and Vaughan (2001) found that inverse-cubed weighting produced the most plausible sur-
 51 face given the data distribution. While IDW is straightforward to implement, computationally efficient, and honours
 52 all measurements, the impact of the power parameter on results can be unintuitive, particularly when dealing with high
 53 variability measurements. That is, lower values of the power parameter may give disproportionately high weighting
 54 to distant points and obscure local trends. In addition, IDW tends to revert produce overly smooth surfaces by biasing
 55 estimates towards the mean, and can not produce interpolated values outside the range of the observed data.

56 A number of alternatives to the IDW method have been proposed. The most ubiquitous is Kriging, a form of
 57 Gaussian Process Regression, which employs spatial correlations across a dataset to generate predictions (Webster and
 58 Oliver, 2007). Comparative testing of the Kriging method was evaluated during the development of both Bedmap2 (Fretwell
 59 et al., 2013) and Bedmap3 (Pritchard et al., 2025), and found to produce less accurate interpolations of Antarctic bed
 60 topography than the adapted spline method used in these well known topography products. Kriging incorporates sta-
 61 tistical, rather than heuristic, estimates of uncertainty and assumes that spatial distributions can be represented as
 62 realisations of random functions. However, its implementation often requires human input, particularly for selecting
 63 model parameters. This means that key decisions, such as choosing variogram models (mathematical functions that
 64 describe how the similarity between data points changes with distance) and addressing gaps in the data where critical
 65 surface features may be missing, can be highly subjective. For example, the interpolation methods used in Bedmap1
 66 and Bedmap2 rely on the assumption that in mountainous areas, where ice fills the valleys, there is a general corre-
 67 lation between ice thickness and the distance from rock outcrops (Fretwell et al., 2013). The problem with such an
 68 assumption is the introduction of systematic biases in ice thickness estimates, particularly in regions with complex
 69 bed topography. These biases can lead to underestimation of bed elevation in deep valleys and overestimation in areas
 70 with irregular subglacial features, affecting subsequent ice volume calculations. Kriging is particularly effective for
 71 modelling phenomena with significant random components (Mitas and Mitasova, 1999) and provides reliable uncer-

72 tainty estimates. However, its performance is highly dependent on expert-driven parameter selection, which must be
 73 carefully adapted to the specific characteristics of Antarctic bed topography data.

74 The key technique used in Bedmap2 is Topogrid, an adapted thin plate spline interpolation method that incorporates
 75 iterative finite difference interpolation to prevent spurious sinks (Fretwell et al., 2013). To determine the optimal
 76 interpolation method to generate Bedmap2, the performance of multiple interpolation methods was compared for
 77 a Scottish Highlands test landscape with sampling patterns mimicking Antarctic surveys. Topogrid outperformed
 78 both spline-with-tension (85 m standard deviation) and Kriging methods, with a standard deviation of 66 m. Despite
 79 this superior performance, spline methods may not honour all data points exactly when producing smooth surfaces.
 80 Furthermore, optimisation of the smoothing parameters remains challenging. As Fretwell et al. noted, “disparities
 81 between varied input data sources, the inhomogeneous spatial distribution of data, and its highly-variable reliability”
 82 necessitated “a rather complicated, multi-stepped process of automatic analyses and manual intervention” (Fretwell
 83 et al., 2013). Their assessment revealed that interpolation errors vary systematically with terrain type (alpine, low relief,
 84 trough, mixed), suggesting that interpolation methods should ideally be tailored to specific landscape morphologies
 85 rather than applying a single method uniformly.

86 Interpolation beyond data points is problematic regardless of the method used. This is particularly challenging
 87 for Antarctica, where data coverage is sparse and irregular. In Bedmap2, 66% of cells have no direct measurements
 88 and rely on extrapolation, with errors increasing with increasing distance from observations at a rate of 2–8 m/km.
 89 Conservative error estimates in interpolated regions reach ± 300 m, with the largest errors (~ 1800 m) occurring when
 90 crossing unsampled troughs in high-relief areas (Fretwell et al., 2013).

91 Mass conservation has been recently developed to address some of the limitations of other methods by incorporat-
 92 ing physics-based constraints into the interpolation process. This approach maintains a consistent balance between the
 93 surface mass balance (accumulation and ablation), ice flux divergence, and spatial thickness changes throughout the
 94 model domain (Roberts et al., 2011). However, these methods require additional input data beyond bed measurements,
 95 including surface velocity fields and assumptions about boundary conditions (e.g. the basal friction coefficient in the
 96 friction law), which may not always be available. Notably, the BedMachine Antarctica and Greenland datasets imple-
 97 ment mass conservation methods and combine radar-derived estimates of ice thickness, gravity-derived bathymetry, ice
 98 velocity, and surface mass balance to create high-resolution gridded bed topography datasets (Morlighem et al., 2020)
 99 which have revealed previously unknown basal features with significant implications for the assessment of glacier vul-
 100 nerability.. The BedMachine algorithm solves a steady hyperbolic partial differential equation that relates the ice flux
 101 divergence and the apparent mass balance (Morlighem et al., 2011, 2020). Here, the apparent mass balance represents
 102 the net local gain or loss of ice mass required to reconcile observed ice thickness and flow.

103 While the above methods have proven useful in many applications, they exhibit several limitations that motivate the

development of new approaches. In particular, many of the above methods lack robust uncertainty estimates, which are crucial for understanding the reliability of interpolated values and planning future measurement campaigns. In addition, these methods often fail to realistically capture the spatial correlation of errors, which can potentially lead to underestimation of uncertainty in regions far from measurements. This highlights the need for new approaches that can: (1) accurately reproduce values at measurement points while providing reasonable extrapolation beyond the data extent; (2) operate with intuitive parameters that users can confidently adjust based on their understanding of the physical system; and (3) provide robust uncertainty estimates that account for measurement errors and interpolation uncertainties. Additionally, modern applications increasingly require methods that can handle large datasets efficiently while maintaining accuracy across various spatial scales.

In this study, we introduce a new ensemble based interpolation algorithm — Stochastic Meshless Uncertainty Gridding (SMUG) — for irregularly sampled data. We use this new algorithm to interpolate the bedrock topography for the Aurora Subglacial Basin, East Antarctic, whilst simultaneously generating robust uncertainty estimates, and compare our bedrock to other recent interpolations. We also evaluate the skill of SMUG in this region by performing additional SMUG interpolations while excluding observational data within various distances from the interpolation point.

2. Methods

Conceptually SMUG is a point-by-point meshing scheme, where each mesh point is considered independently, and an interpolated value returned if there are no observations at the mesh point. SMUG uses a pool of neighbouring observation points (described in section 2.1) and a Taylor series approximation (described in section 2.2) to estimate the value, and uncertainty, of the data field being interpolated at each mesh point. An ensemble of estimates are calculated, and the inverse-uncertainty-squared weighted central tendency and spread of the resulting ensemble returned. The derivatives required for the Taylor series approximation are calculated using the meshless method of Roberts (2024), which ensures machine-precision accuracy for randomly located neighbours (compared with most methods for estimating derivatives which require regularly-distributed data for high accuracy).

2.1. Neighbour selection

Neighbour selection is a multi-step process. Firstly, a pool of suitable neighbours representing broad angular coverage is determined. Then the actual points used for the Taylor series approximation for each ensemble member is selected from within this pool.

132 **2.1.1. Neighbour pool**

133 The selection of neighbouring observation points to include in the pool of suitable points is via an angular-
 134 separation weighted inverse distance greedy algorithm. First, the input data is divided into equi-angular sectors, rotated
 135 such that the closest observation to the interpolation point of interest is in the centre of the first sector. By default,
 136 there are eight sectors for a first-order Taylor series approximation and 16 for a second-order approximation, but the
 137 user can override these defaults. The pool of suitable neighbours will contain a default of four points per sector (again
 138 the user can override this default). Second, on a sector-by-sector basis, the closest point is included, and the effective
 139 distance to each remaining point is scaled by a factor d_{eff} to penalise points close (in an angular sense) to the points
 140 in the sector already selected. Here,

$$d_{\text{eff}} = (1 + \exp(-(2\delta\theta/\Theta)^2)), \quad (1)$$

141 where Θ is the angular extent of each sector and $\delta\theta$ the angular separation between the point under consideration
 142 and the previous point added to the pool from this sector. The next closest point (based on the penalised distance) is
 143 selected and the process repeated until the required number of points have been added to the pool from each sector in
 144 turn.

145 **2.1.2. Ensemble member point selection**

146 Each ensemble member uses its own subset of points of N points drawn from the pool of suitable points (sec-
 147 tion 2.1.1). By default $N=4$ for a first order Taylor series approximation, and 8 for a second order approximation,
 148 but again these defaults can be overridden. First, if N is less than the number of sectors, we uniform-randomly select
 149 which sectors to draw a single point from, always including the sector containing the closest point to our point of
 150 interest. Second, from each chosen sector, we non-uniformly select which point from the pool to include, weighted
 151 by the inverse (true, not effective) distance squared to the point. If N is larger than the number of sectors, we then
 152 inverse-distance-squared weighted randomly select from the remaining points in the pool.

153 Finally, we again inverse-distance-squared weighted randomly choose which of the N points from which to perform
 154 the Taylor series approximation.

155 **2.2. SMUG**

156 The SMUG algorithm is based around a Taylor series approximation (of order 1 or 2, as determined by the user)
 157 from a subset of size N points randomly drawn from the pool of suitable neighbouring points (section 2.1) to estimate
 158 the value of the observational data field of interest, and its associated uncertainty, at the mesh point (P) of interest.

159 The Taylor series approximation is conducted from one of the N neighbours (N_T), chosen randomly with an inverse
 160 distance (to P) squared weighting. To simplify the calculations, the N neighbouring points are translated and rotated
 161 such that P is at the origin, and N_T is on the positive x-axis. The Taylor series estimate (T) is then given by:

$$T = \begin{cases} V_T - V'_T * X_T, & \text{for order 1;} \\ V_T - V'_T * X_T + V''_T * X_T^2 / 2, & \text{for order 2;} \end{cases} \quad (2)$$

162 Here, X_T is the x-axis intercept of the rotated point N_T , V_T is the observation data field of interest at the point
 163 X_T , and V'_T and V''_T the first and second order partial derivative, respectively, of the observational data field of interest
 164 calculated in the x direction at the point X_T . These derivatives are estimated using a meshless method based on
 165 differential quadrature, radial basis test functions and hyper-dual numbers (Roberts, 2024). The meshless method
 166 ingests the locations of the neighbouring observational points and calculates a corresponding vector of weighting
 167 coefficients to be applied to the observational points to estimate the required derivatives. When implementing the
 168 projection of the gradient given in Equation 2 we use a multi-step approach so that the distance X_T is subdivided
 169 into m intervals. This multi-stage approach is beneficial when the interpolation distances are large and the gradient
 170 may vary significantly, and improves both the estimate and uncertainty. The derivative is estimated at the start of the
 171 interval, projected along the length $\frac{X_T}{m}$ and the process repeated until the interpolation point is reached, taking into
 172 account changes in the gradient. This increases the effective order of the Taylor series expansion such that it is higher
 173 than would be the case for a single step approach.

174 For the Taylor expansion around N_T , the uncertainty is estimated using Ball-Arithmetic (Johansson, 2020). Specif-
 175 ically, we apply the following three rules to the Taylor series estimates of Equation 2:

$$\begin{aligned} (A_c, A_r) + (B_c, B_r) &= (A_c + B_c, A_r + B_r) && \text{addition} \\ (A_c, A_r) - (B_c, B_r) &= (A_c - B_c, A_r + B_r) && \text{subtraction} \\ (A_c, A_r) \times (B_c, B_r) &= (A_c \times B_c, |A_c|B_r + |B_c|A_r + A_r B_r) && \text{multiplication} \end{aligned} \quad (3)$$

176 Here, two floating point numbers (A and B) are each represented by a pair of numbers (c, r) , where c is the centre (best
 177 approximation), and r is the (always non-negative) radius (uncertainty) of the floating point ball approximation of a
 178 real number.

179 To obtain robust estimates of the observational data field of interest and uncertainty at P , we generate an ensem-
 180 ble of estimates and calculate the uncertainty-weighted central tendency and spread using robust statistical measures

181 (M-estimator for central tendency (Wilcox, 2010) and FQn for spread (Smirnov and Shevlyakov, 2014)): these are
 182 statistically more robust equivalents of the mean and standard deviation respectively. Specifically, we generate an
 183 ensemble of 500 (default, but user changeable) members using random (but not uniformly) selected points, and calcu-
 184 late the inverse-uncertainty-squared weighted central tendency (M-estimator) and spread (FQn) calibrated to give the
 185 standard deviation for normally-distributed data.

186 **2.3. Application to East Antarctic Bed Topography**

187 We test the SMUG algorithm using the Bedmap3 observational datasets (Frémand et al., 2023) of bed elevation in
 188 a domain of East Antarctica that extends from the West Ice Shelf to Porpoise Bay to the east (Figure 1). The inland
 189 limit of the region includes Lake Vostok and Dome C; hence, our domain encompasses the Aurora Subglacial Basin
 190 (ASB) and the Denman-Scott catchment, the key outlet glaciers of which (Totten and Denman Glaciers) are among
 191 the most rapidly thinning glaciers in East Antarctica (Smith et al., 2020).

192 The Bedmap3 dataset is collated from radar surveys of Antarctica beginning in the 1950s; here, we apply the
 193 SMUG algorithm to the observational datasets included in the Bedmap3 compilation, which are freely accessible via
 194 the UK Polar Data Centre and the SCAR Bedmap Data Portal Frémand et al. (2023). All original data are provided
 195 on a WSG84 Latitude-Longitude grid (EPSG:4326), which we then convert to Antarctic Polar Stereographic projec-
 196 tion (EPSG:3031), referenced to a Standard Parallel of 71°S. We conduct a basic quality check of the bed elevation
 197 data, removing several outlier data lines (where bed elevation measurements at cross-over points are inconsistent with
 198 the remaining data), specifically lines 18271–19771 from the STANFORD_1971_SPRI-NSF-TUD_AIR_BM3.csv and
 199 segment 2012331_ASB_JKB2h_R11Ec from UTIG_2010_ICECAP_AIR_BM3.csv.

200 We apply the SMUG algorithm described in sections 2.1 and 2.2 above to our domain of interest, to derive a gridded
 201 bed elevation reconstruction at 500 m, and defined on the same grid as the BedMachine Antarctica version 2 dataset
 202 (Morlighem et al., 2020) for comparison with both BedMachine and Bedmap3 gridded bedrock elevation products
 203 (Pritchard et al., 2024) .

204 **3. Results**

205 **3.1. SMUG bed elevation and uncertainty**

206 A bed elevation reconstruction for our domain of interest using a Taylor series expansion of order of 1 is shown
 207 in Figure 1A. The SMUG interpolation captures the known features of the region consistent with other bed elevation
 208 products. Much of the ASB bed is below sea level extending far inland as well as features above sea level, namely
 209 Highlands A and B, deep troughs of the Vanderford Subglacial Trench separating Law Dome from the main Antarctic
 210 continent and flowing into the Totten and Vanderford glaciers. To lessen the effect of artifacts such noise occurring in

Table 1

Summary of differences from published bed topographies.

	mean difference (m)	median difference (m)	standard deviation (m)	Inter-quartile range(m)
SMUG-bedmap3	0.79	-14.64	163.93	87.68
SMUG-bedmachine	4.48	5.87	178.62	115.35

211 areas of rapidly changing terrain, we applied a 1km median filter over the bedrock elevation and extracted the locations
 212 of highest and deepest points for comparison with other bedrock elevation products. The highest point then estimated
 213 by SMUG is at 2081 ± 55 m, in the Vostok subglacial highlands to the south-west of Lake Vostok. Other high elevation
 214 regions, such as Highlands A and B and Dome C also resolved. The lowest estimated bedrock elevation regions occurs
 215 at the trough at the front of the Denman Glacier at -2833 ± 66 m. The deep bedrock of the Vanderford Subglacial
 216 Trench, near the Moscow University Ice Shelf and deep channels inland in the Aurora Subglacial Basin are also well
 217 resolved.

218 The uncertainties in the bedrock elevation, shown in Figure 1B, are highest where there greater spacing between
 219 observational data, such as the large regions to the north-east of Dome C and south-qast of Lake Vostok. Large errors
 220 are also apparent where there are rapid changes in elevation over short distances, such as along the boundaries of the
 221 deep channels in the Denman-Scott region.

222 Figure 1C–D shows the difference between the SMUG-derived bedrock elevation and that of both Bedmap3
 223 (Pritchard et al. (2025)) and Bedmachine (Morlighem et al. (2020)), noting that except for the exclusion of the data
 224 outliers listed above, SMUG uses the same observational dataset as Bedmap3.

225 Overall, SMUG is generally in good agreement with both Bedmap3 and BedMachine, with a small bias and moder-
 226 ate differences (see Table 1). However, SMUG predicts a generally shallower topography, and larger local differences,
 227 particularly in regions where we have sparse data and rough topography, where SMUG uncertainty is higher and all
 228 three methods are limited by data availability.

229 The Denman-Scott region in particular shows marked differences between SMUG and other bed topography inter-
 230 polation products. There is a good coverage of radar data in parts of the region, with several surveys conducted over
 231 the region with line spacing of 5 km perpendicular to the flow direction at the front of the Denman glacier, increasing
 232 to 10 km further inland as well as flight lines along the direction of flow. SMUG estimates the Denman Glacier trough
 233 as generally more shallow than both Bedmap3 and Bedmachine. The steep gradient in deeper parts of this trough
 234 results in missing observational data where radar reflections are not returned. In particular, the transect from Ta–Tb
 235 (yellow line in Figure 2A–F) which is along a flight line which loses radar reflections across the deeper parts of the
 236 trough, is considerably more shallow with the SMUG estimate differing by approximately 2000 m from Bedmap3 and
 237 Bedmachine along this transect .

238 To further investigate this difference, shown in Figure 2G, we compared the bed with the MeASURES surface
 239 speed and computed the longitudinal strain rate from the MeASURES surface velocity in the Denman region, shown
 240 in Figure 2E–F. We see that the flow speed at the location of the Ta–Tb transect is slightly slower than directly up and
 241 downstream and suggests a slight shoaling from a shallower depth at that point. Further evidence of this is seen in the
 242 pattern of compression and extension in the longitudinal strain rate at the same location.

243 A grid pattern is visible in the Lake Vostok bathymetry, aligning with the 1000 m below mean sea level contour.
 244 In this region, there are at least two observational survey datasets, which appear to have a vertical offset between their
 245 observations of bed elevation. The more uniformly spaced survey reports, in general, slightly shallower bed elevations
 246 compared to the other survey which has a higher density sampling in the along track, but has larger spacing between
 247 survey tracks. This results in interpolated topography that is in general slightly shallower than the -1000 m contour
 248 except near the locations of the more spatially anisotropic survey.

249 3.2. Skill and bias

250 To investigate the skill and bias of the SMUG algorithm and the robustness of the uncertainty estimates in regions
 251 of sparse or non-existent data, we generated a series of 1 km resolution interpolations of our domain, which excluded
 252 any observational data within a series of distances to the interpolation point, namely at distances of 5, 10, 20, 30, 40
 253 and 50 km. The results of this data exclusion exercise are shown in Figure 3, with spatial maps for individual data
 254 exclusion distances shown in Figure 4. In each case we compare the resultant interpolation with that using all data.

255 The order of the Taylor series expansion used in SMUG is its main user selectable parameter. To demonstrate the
 256 effect of increasing the order of the Taylor series, we also performed a 2nd order Taylor interpolation for the ASB
 257 region, shown in Figure 5. Compared to the 1st order Taylor interpolation (Figure 1), the higher order interpolation
 258 has more impulse noise and higher uncertainty, but consistent large scale features. There is low bias between the two
 259 interpolations (mean difference of 2.2 m and a median of 0.0 m) and, in general, low differences (standard deviation
 260 of 115.6 m and an inter-quartile range of 57.4 m). The differences tend to be lower across the deep basins where the
 261 bedrock is below mean sea-level, and large over the higher elevation features, although these also tend to be the areas
 262 with larger distances between survey lines.

263 4. Discussion

264 SMUG is a stochastic algorithm, using random numbers to select which combinations of points from the pool of
 265 nearby observations that are used in constructing the ensemble. Therefore, the interpolated data-field (and associated
 266 uncertainties) will differ slightly between successive applications to the same dataset, except at any points where
 267 observations exist. These differences should be small with an appropriately chosen ensemble size (typically around 500

268 ensemble members is sufficient, but may be increased for datasets with particularly challenging distributions).
 269 However, larger ensemble sizes require more computational resources, with the time taken scaling approximately
 270 linearly with the ensemble size. The majority of SMUG's run-time (with the exact proportion varying with the number
 271 of neighbours and their spatial distribution) is taken to calculate the spatial derivatives, and scales linearly with the
 272 number of ensemble members. The routines to calculate the ensemble statistics (M-estimator and FQn) is only a minor
 273 contribution, and scales non-linearly with the number of ensemble members.

274 When interpolating a point, the selection of observations included can have significant impact on the quality and
 275 fidelity of the resulting interpolation. No single interpolation scheme is likely to perform well for all possible combi-
 276 nations of underlying field and observational sampling, and can lead to artefacts in the resulting interpolated field. We
 277 believe that the scheme used in SMUG — a pool of nearby points with good angular coverage — combined with the
 278 stochastic ensemble result in a reasonable and skilful interpolation. However, as SMUG considers each interpolation
 279 point in isolation, if the in-built point selection routine does not perform well in certain situations, the user can easily
 280 bypass the built-in routine by, for each interpolation point, pre-processing the observational data to select N points
 281 using whatever method they so choose.

282 The SMUG interpolation value is calculated by projecting the estimated local gradient. As such, it will perform
 283 better in situations where that gradient is smooth, and the distribution of observations allow for it to be accurately
 284 estimated. Therefore, we expect SMUG to perform better in situations where the underlying data field is smoothly
 285 varying and sampled sufficiently densely to capture changes in the gradient of that field. When the underlying data
 286 field is more uneven, and sparsely sampled, SMUG's random sampling of the proximal observations, and relatively
 287 large ensemble size should help minimise any bias, and produce a robust estimation of the uncertainty.

288 As SMUG considers each interpolation point in isolation, impulse noise (also known as "salt and pepper noise")
 289 can occur, although this should be minimal with an appropriate ensemble size. This effect is present in our Antarctic
 290 bedrock topography example, and is seen to be present in the inland region near south-west of Lake Vostok, and at
 291 the coastal margin between 80 – 85°E. If such noise is present, it can easily be removed using a median filter (or
 292 similar), although care should be taken to honour observational data points when filtering. The resultant topography
 293 after applying the median filter is shown in Fig. 6.

294 As the exclusion distance increases, meaning less data is used, the ability to resolve features decreases (and un-
 295 certainty increases). This is seen in Figure 4, where features such as the inland Highlands are no longer able to be
 296 resolved beyond an exclusion radius of 20km, and the deep subglacial troughs past a 10km exclusion radius. When
 297 comparing the interpolation with the full observational dataset to the cases where data is excluded, SMUG shows very
 298 low biases and low differences (with a median difference of less than 150 m when data is excluded within 50 km). As
 299 SMUG uses a Taylor expansion, and therefore projects observed gradients, it is not surprising that it cannot resolve

300 small scale (compared to the data exclusion radius) features in the absence of data. Additional data-streams, such as
 301 surface elevation and velocities as used by mass conservation based interpolation schemes (such as BedMachine), are
 302 required to resolve features in the absence of proximal direct observations.

303 For Antarctic bedrock interpolation, mass conserving schemes use the local velocity field as an input data-stream.
 304 This can result in anisotropic interpolation performance between the along streamline and across flow directions (e.g.
 305 Roberts et al., 2011). As SMUG only uses observed bedrock elevations, it does not suffer from streamline anisotropy,
 306 but may suffer from anisotropic effects due to the distribution of observations, such as the region shown in Fig. 7.

307 In addition to producing interpolations, SMUG also calculates the associated uncertainties using a combination of
 308 contributions from the ensemble (using different points for both the derivative calculation and the origin of the Tay-
 309 lor series expansion), and the uncertainty in the observations. In the context of producing and utilising an Antarctic
 310 bedrock estimation, these uncertainties have several direct applications: (1) quantifying where we are the most uncer-
 311 tain about the bed topography, given the current distribution of the observation; (2) as input data for ice sheet models to
 312 assess where these uncertainties have the greatest impact on ice dynamics, and the associated Antarctic contribution to
 313 local and global sea-level change; (3) using data from both of the previous applications, designing future observational
 314 campaigns to meaningfully reduce uncertainties around Antarctica's sea-level contributions; and (4) for comparison
 315 with other interpolation algorithms.

316 Robust interpolation uncertainty estimates (such as those produced by SMUG), are still subject to errors, biases
 317 and limitations associated with the underlying algorithm, its assumptions, and the distribution, errors and uncertainties
 318 in the observations. One potentially useful approach to reduce and/or quantify these errors and biases is to produce
 319 interpolations using a multi-model-ensemble built on different algorithms, with the final interpolation being based
 320 on robust statistical measures of the central tendency and spread of the multi-model-ensemble (see, for example in
 321 an process optimisation context Liu, 2018). With the use of robust statistical measures (such as the median or M-
 322 estimator), such multi-model-ensembles are largely unaffected by outliers. In addition, the individual model estimated
 323 uncertainties can be used to inverse-weight the final interpolation to de-emphasise models with large uncertainties.

324 The main user-configurable parameter for SMUG is the order of the Taylor series expansion. As for polynomial
 325 interpolation, in general using the lowest order interpolation that is consistent with the data will produce the best
 326 results, i.e. typically first order SMUG interpolation is appropriate, although in some circumstances second order
 327 might perform better. At high observational density (and associated small interpolation distances), the differences
 328 between first and second order interpolations should be small. In fact, even at the (locally) relatively large observational
 329 data spacing over the Aurora Subglacial Basin, the differences in the interpolated bed topography using both first
 330 and second order interpolation are small (compare Figures 1 and 5). Due to the additional terms in the Taylor series
 331 expansion (Equation 2), the generally larger weighting coefficients associated with the meshless derivative calculations

332 at higher order (e.g. Abramowitz and Stegun, 1968; Roberts, 2008), and the monotonically increasing nature of the
333 ball-arithmetic uncertainty (Equation 3), the uncertainty estimates for second order interpolations will be larger than
334 for first order.

335 5. Conclusions

336 We present a new method for interpolating sparse and unevenly sampled spatial data with robust uncertainty es-
337 timates, namely the Stochastic Meshless Uncertainty Gridding (SMUG) algorithm. To demonstrate its performance,
338 we apply SMUG to reconstruct the bed topography of the Aurora Subglacial Basin in East Antarctica. As with all
339 interpolation approaches, SMUG is subject to artefacts arising from the quality of input data and can introduce im-
340 pulse noise in the resulting interpolated data field. While such issues can be partly mitigated through quality control
341 of the input data and varying the ensemble size, challenges will persist where observational coverage is limited. A key
342 strength of SMUG lies in its ability to produce robust, ensemble-based uncertainty estimates. Future work could ex-
343 plore extensions such as meta-ensembles that combine multiple algorithms to further improve accuracy and uncertainty
344 characterisation.

345 6. Acknowledgements

346 The authors would like to acknowledge ...

347 **Code availability section**

348 FORTRAN source code for SMUG is available for download under the GPLv3 licence at <https://github.com/jlr581/SMUG>.
 349 The meshless library code SMUG uses is included within this repository, but is also available at
 350 https://github.com/jlr581/finite_difference.

351 **A. Supplementary figures**352 **References**

- 353 Abramowitz, M., Stegun, I. (Eds.), 1968. Handbook of Mathematical Functions. United States Department of Commerce, Washington.
- 354 Frémand, A.C., Fretwell, P., Bodart, J.A., Pritchard, H.D., Aitken, A., Bamber, J.L., Bell, R., Bianchi, C., Bingham, R.G., Blankenship, D.D.,
 355 Casassa, G., Catania, G., Christianson, K., Conway, H., Corr, H.F.J., Cui, X., Damaske, D., Damm, V., Drews, R., Eagles, G., Eisen, O.,
 356 Eisermann, H., Ferraccioli, F., Field, E., Forsberg, R., Franke, S., Fujita, S., Gim, Y., Goel, V., Gogineni, S.P., Greenbaum, J., Hills, B.,
 357 Hindmarsh, R.C.A., Hoffman, A.O., Holmlund, P., Holschuh, N., Holt, J.W., Horlings, A.N., Humbert, A., Jacobel, R.W., Jansen, D., Jenkins,
 358 A., Jokat, W., Jordan, T., King, E., Kohler, J., Krabill, W., Kusk Gillespie, M., Langley, K., Lee, J., Leitchenkov, G., Leuschen, C., Luyendyk, B.,
 359 MacGregor, J., MacKie, E., Matsuoka, K., Morlighem, M., Mouginot, J., Nitsche, F.O., Nogi, Y., Nost, O.A., Paden, J., Pattyn, F., Popov, S.V.,
 360 Rignot, E., Rippin, D.M., Rivera, A., Roberts, J., Ross, N., Ruppel, A., Schroeder, D.M., Siegert, M.J., Smith, A.M., Steinhage, D., Studinger,
 361 M., Sun, B., Tabacco, I., Tinto, K., Urbini, S., Vaughan, D., Welch, B.C., Wilson, D.S., Young, D.A., Zirizzotti, A., 2023. Antarctic Bedmap
 362 data: Findable, Accessible, Interoperable, and Reusable (FAIR) sharing of 60 years of ice bed, surface, and thickness data. Earth System Science
 363 Data 15, 2695–2710. doi:10.5194/essd-15-2695-2023.
- 364 Fretwell, P., Pritchard, H.D., Vaughan, D.G., Bamber, J.L., Barrand, N.E., Bell, R., Bianchi, C., Bingham, R.G., Blankenship, D.D., Casassa,
 365 G., Catania, G., Callens, D., Conway, H., Cook, A.J., Corr, H.F.J., Damaske, D., Damm, V., Ferraccioli, F., Forsberg, R., Fujita, S., Gim, Y.,
 366 Gogineni, P., Griggs, J.A., Hindmarsh, R.C.A., Holmlund, P., Holt, J.W., Jacobel, R.W., Jenkins, A., Jokat, W., Jordan, T., King, E.C., Kohler,
 367 J., Krabill, W., Riger-Kusk, M., Langley, K.A., Leitchenkov, G., Leuschen, C., Luyendyk, B.P., Matsuoka, K., Mouginot, J., Nitsche, F.O., Nogi,
 368 Y., Nost, O.A., Popov, S.V., Rignot, E., Rippin, D.M., Rivera, A., Roberts, J., Ross, N., Siegert, M.J., Smith, A.M., Steinhage, D., Studinger,
 369 M., Sun, B., Tinto, B.K., Welch, B.C., Wilson, D., Young, D.A., Xiangbin, C., Zirizzotti, A., 2013. Bedmap2: improved ice bed, surface and
 370 thickness datasets for Antarctica. The Cryosphere 7, 375–393. doi:10.5194/tc-7-375-2013.
- 371 Graham, F.S., Roberts, J.L., Galton-Fenzi, B.K., Young, D., Blankenship, D., Siegert, M.J., 2017. A high-resolution synthetic bed elevation grid of
 372 the Antarctic continent. Earth System Science Data 9, 267–279. doi:10.5194/essd-9-267-2017.
- 373 Johansson, F., 2020. Ball arithmetic as a tool in computer algebra, in: Gerhard, J., Kotsireas, I. (Eds.), Maple in Mathematics Education and
 374 Research, Springer International Publishing, Cham. pp. 334–336.
- 375 Liu, K.N., 2018. A multi-model ensemble approach to process optimization considering model uncertainty. Journal of Industrial and Production
 376 Engineering 35, 550–557. doi:10.1080/21681015.2018.1534756.
- 377 Lythe, M.B., Vaughan, D.G., 2001. BEDMAP: A new ice thickness and subglacial topographic model of Antarctica. Journal of Geophysical
 378 Research: Solid Earth 106, 11335–11351. URL: <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2000JB900449>,
 379 doi:<https://doi.org/10.1029/2000JB900449>.
- 380 MacKie, E.J., Schroeder, D.M., 2020. Geostatistically simulating subglacial topography with synthetic training data, in: IGARSS 2020 - 2020 IEEE
 381 International Geoscience and Remote Sensing Symposium, pp. 2991–2994. doi:10.1109/IGARSS39084.2020.9324563.

Stochastic Meshless Uncertainty Gridding

- 382 MacKie, E.J., Schroeder, D.M., Zuo, C., Yin, Z., Caers, J., 2021. Stochastic modeling of subglacial topography exposes uncertainty in water routing
383 at Jakobshavn Glacier. *Journal of Glaciology* 67, 75–83. doi:10.1017/jog.2020.84.
- 384 McCormack, F.S., Stål, T., Shao, N., MacKie, E., Hinojosa, A.F., Lösing, M., Roberts, J.L., Ehrenfeucht, S., Dow, C.F., Submitted. Synthetic bed
385 topographies for Antarctica and their utility in ice sheet modelling. *Proceedings of the Royal Society A* .
- 386 Mitas, L., Mitasova, H., 1999. Spatial Interpolation. *GeoInformation International*, Wiley. chapter 34. pp. 481–492.
- 387 Morlighem, M., Rignot, E., Binder, T., Blankenship, D., Drews, R., Eagles, G., Eisen, O., Ferraccioli, F., Forsberg, R., Fretwell, P., Goel, V.,
388 Greenbaum, J.S., Gudmundsson, H., Guo, J., Helm, V., Hofstede, C., Howat, I., Humbert, A., Jokat, W., Karlsson, N.B., Lee, W.L., Matsuoka,
389 K., Millan, R., Mouginot, J., Paden, J., Pattyn, F., Roberts, J., Rosier, S., Ruppel, A., Seroussi, H., Smith, E.C., Steinhage, D., Sun, B., van den
390 Broeke, M.R., van Ommen, T.D., van Wessem, M., Young, D.A., 2020. Deep glacial troughs and stabilizing ridges unveiled beneath the margins
391 of the Antarctic ice sheet. *Nature Geoscience* 13, 132–137. doi:10.1038/s41561-019-0510-8.
- 392 Morlighem, M., Rignot, E., Seroussi, H., Larour, E., Ben Dhia, H., Aubry, D., 2011. A mass conservation approach for mapping glacier ice
393 thickness. *Geophysical Research Letters* 38. URL: <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2011GL048659>,
394 doi:<https://doi.org/10.1029/2011GL048659>.
- 395 Pritchard, H., Fretwell, P., Fremand, A., Bodart, J., Kirkham, J., Aitken, A., Bamber, J., Bell, R., Bianchi, C., Bingham, R., Blankenship, D.,
396 Casassa, G., Catania, G., Christianson, K., Conway, H., Corr, H., Cui, X., Damaske, D., Damn, V., Dorschel, B., Drews, R., Eagles, G., Eisen,
397 O., Eisermann, H., Ferraccioli, F., Field, E., Forsberg, R., Franke, S., Fujita, S., Gim, Y., Goel, V., Gogineni, S.P., Greenbaum, J., Hills, B.,
398 Hindmarsh, R., Hoffman, A., Holmlund, P., Holschuh, N., Holt, J., Horlings, A., Humbert, A., Jacobel, R., Jansen, D., Jenkins, A., Jokat, W.,
399 Jordan, T., King, E., Kohler, J., Krabill, W., Langley, K., Lee, J., Leitchenkov, G., Leuschen, C., Luyendyk, B., MacGregor, J., MacKie, E.,
400 Matsuoka, K., Moholdt, G., Morlighem, M., Mouginot, J., Nitsche, F., Nogi, Y., Nost, O., Paden, J., Pattyn, F., Popov, S., Kusk-Gillespie, M.,
401 Rignot, E., Rippin, D., Rivera, A., Roberts, J., Ross, N., Ruppel, A., Schroeder, D., Siegert, M., Smith, A., Steinhage, D., Studinger, M., Sun,
402 B., Tabacco, I., Tinto, K., Urbini, S., Vaughan, D., Welch, B., Wilson, D., Young, D., Zirizzotti, A., 2024. BEDMAP3 - Ice thickness, bed and
403 surface elevation for Antarctica - gridding products. URL: <https://data.bas.ac.uk/full-record.php?id=GB/NERC/BAS/PDC/01615>,
404 doi:10.5285/2D0E4791-8E20-46A3-80E4-F5F6716025D2.
- 405 Pritchard, H.D., Fretwell, P.T., Fremand, A.C., Bodart, J.A., Kirkham, J.D., Aitken, A., Bamber, J., Bell, R., Bianchi, C., Bingham, R.G., Blanken-
406 ship, D.D., Casassa, G., Christianson, K., Conway, H., Corr, H.F.J., Cui, X., Damaske, D., Damm, V., Dorschel, B., Drews, R., Eagles, G.,
407 Eisen, O., Eisermann, H., Ferraccioli, F., Field, E., Forsberg, R., Franke, S., Goel, V., Gogineni, S.P., Greenbaum, J., Hills, B., Hindmarsh,
408 R.C.A., Hoffman, A.O., Holschuh, N., Holt, J.W., Humbert, A., Jacobel, R.W., Jansen, D., Jenkins, A., Jokat, W., Jong, L., Jordan, T.A., King,
409 E.C., Kohler, J., Krabill, W., Maton, J., Gillespie, M.K., Langley, K., Lee, J., Leitchenkov, G., Leuschen, C., Luyendyk, B., MacGregor, J.A.,
410 MacKie, E., Moholdt, G., Matsuoka, K., Morlighem, M., Mouginot, J., Nitsche, F.O., Nost, O.A., Paden, J., Pattyn, F., Popov, S., Rignot,
411 E., Rippin, D.M., Rivera, A., Roberts, J.L., Ross, N., Ruppel, A., Schroeder, D.M., Siegert, M.J., Smith, A.M., Steinhage, D., Studinger, M.,
412 Sun, B., Tabacco, I., Tinto, K.J., Urbini, S., Vaughan, D.G., Wilson, D.S., Young, D.A., Zirizzotti, A., 2025. Bedmap3 updated ice bed, sur-
413 face and thickness gridded datasets for Antarctica. *Scientific Data* 12, 414. URL: <https://doi.org/10.1038/s41597-025-04672-y>,
414 doi:10.1038/s41597-025-04672-y.
- 415 Roberts, J., 2008. A method for calculating meshless finite difference weights. *International Journal for Numerical Methods in Engineering* 74,
416 321–336. doi:10.1002/nme.2169.
- 417 Roberts, J.L., 2024. Meshless weighting coefficients for arbitrary nodes: The efficient computation to machine precision using hyper-dual numbers.
418 *Advances in Engineering Software* 197, 103753. doi:10.1016/j.advengsoft.2024.103753.
- 419 Roberts, J.L., Warner, R.C., Young, D., Wright, A., van Ommen, T.D., Blankenship, D.D., Siegert, M., Young, N.W., Tabacco, I.E., Forieri, A.,

Stochastic Meshless Uncertainty Gridding

- 420 Passerini, A., Zirizzotti, A., Frezzotti, M., 2011. Refined broad-scale sub-glacial morphology of Aurora Subglacial Basin, East Antarctica derived
421 by an ice-dynamics-based interpolation scheme. *The Cryosphere* 5, 551–560. doi:10.5194/tc-5-551-2011.
- 422 Smirnov, P.O., Shevlyakov, G.L., 2014. Fast highly efficient and robust one-step M-estimators of scale based on Qn. *Computational Statistics &*
423 *Data Analysis* 78, 153–158. doi:10.1016/j.csda.2014.04.013.
- 424 Smith, B., Fricker, H.A., Gardner, A.S., Medley, B., Nilsson, J., Paolo, F.S., Holschuh, N., Adusumilli, S., Brunt, K., Csatho, B., Harbeck, K.,
425 Markus, T., Neumann, T., Siegfried, M.R., Zwally, H.J., 2020. Pervasive ice sheet mass loss reflects competing ocean and atmosphere processes.
426 *Science* 368, 1239–1242. doi:10.1126/science.aaz5845.
- 427 Webster, R., Oliver, M., 2007. *Geostatistics for Environmental Scientists*. Wiley. doi:10.1002/9780470517277.
- 428 Wilcox, R.R., 2010. *Fundamentals of Modern Statistical Methods*. Springer-Verlag, New York. doi:10.1007/978-1-4419-5525-8.

429 **List of Figures**

430 1	Antarctic bed topography reconstruction at 500m horizontal resolution, and differences from published interpolations. A) SMUG interpolation, with 500m elevation contours shown in grey. B) Interpolation uncertainty and observational data locations (blue points). C) SMUG minus Bedmap3 Pritchard et al. (2025). D) SMUG minus Bedmachine V2 Morlighem et al. (2020)	17
431 2	Denman Glacier local region. A) SMUG interpolation, B) Bedmap3 interpolation, C) Bedmachine interpolation, D) Bedmap3 observational data, E) MeASUREs surface speed, F) longitudinal strain rate from MeASUREs surface velocity, G) transect Ta–Tb showing SMUG (gray shading shows un- certainty), Bedmap3, BedMachine bed profiles, and the observations.	18
432 3	Impact of excluding local data on the interpolation compared to no data exclusion at 1km resolution. Shown are the mean difference (circle), \pm one standard deviation (error bars) and lower quartile, me- dian and upper quartile (boxes).	19
433 4	Antarctic bed topography (first and third rows) at 1km horizontal resolution, with observational data excluded, and corresponding difference from interpolation without data exclusion (second and forth rows). Data exclusion radius 5 km (A and D), 10 km (B and E), 20 km (C and F), 30 km, (G and J), 40 km (H and K) and 50 km (I and L).	20
434 5	2nd order Taylor interpolation of ASB at 1 km reconstruction, raw interpolation. A) Bedrock topog- raphy, and B) interpolation uncertainty (grey-scale) and observational data locations (blue points). . .	21
435 6	The bed topography dataset after a median filter has been applied to remove impulse noise. The blue star indicates the location of the deepest topography and the red star the highest point in the bedrock. .	22
436 7	Example of anisotropic effects due to distribution of observations, noting the striping present and dis- continuity where there are fewer observations.	23
437		
438		
439		
440		
441		
442		
443		
444		
445		
446		
447		
448		
449		
450		

Stochastic Meshless Uncertainty Gridding

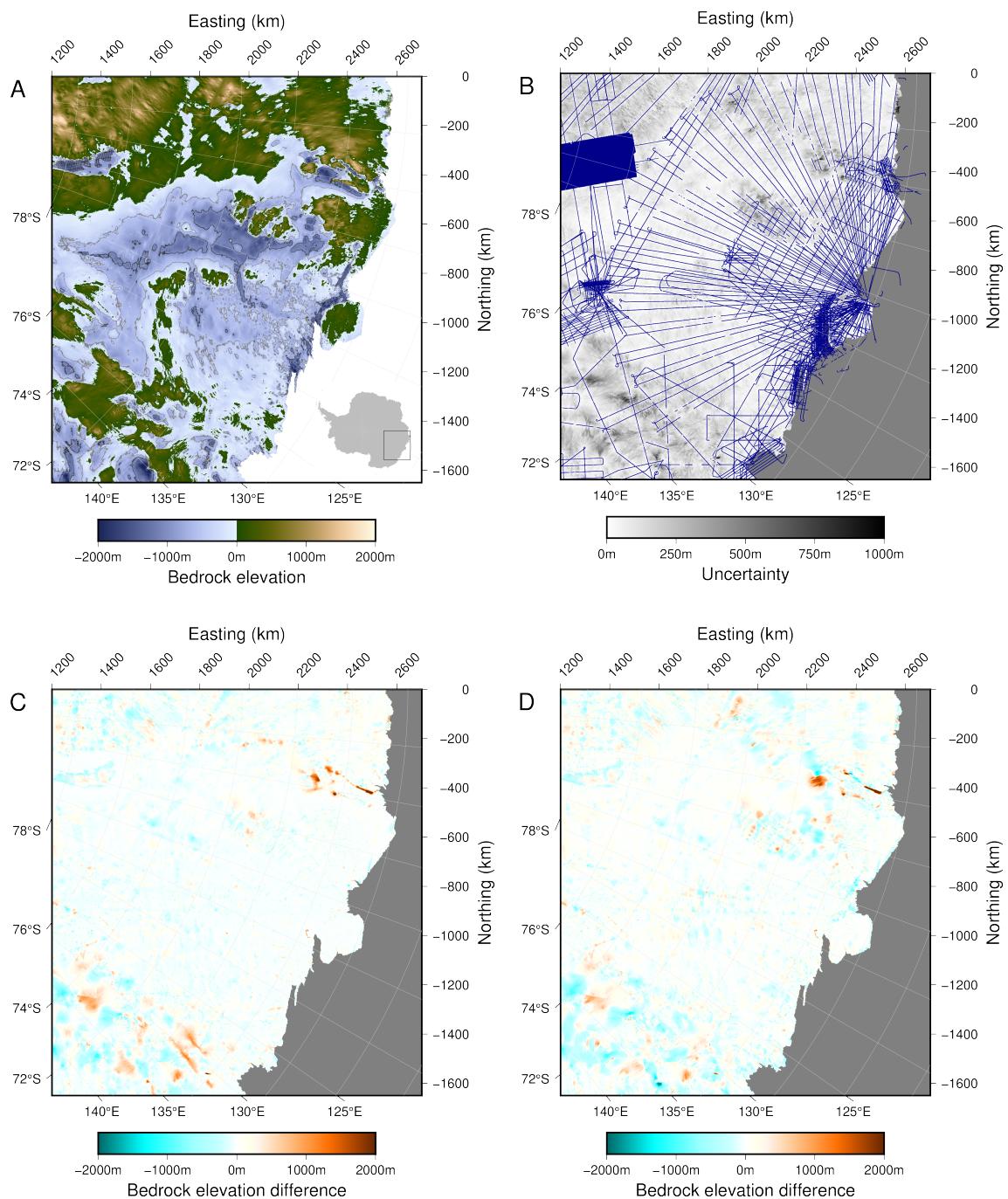


Figure 1: Antarctic bed topography reconstruction at 500m horizontal resolution, and differences from published interpolations. A) SMUG interpolation, with 500m elevation contours shown in grey. B) Interpolation uncertainty and observational data locations (blue points). C) SMUG minus Bedmap3 Pritchard et al. (2025). D) SMUG minus BedMachine V2 Morlighem et al. (2020)

Stochastic Meshless Uncertainty Gridding

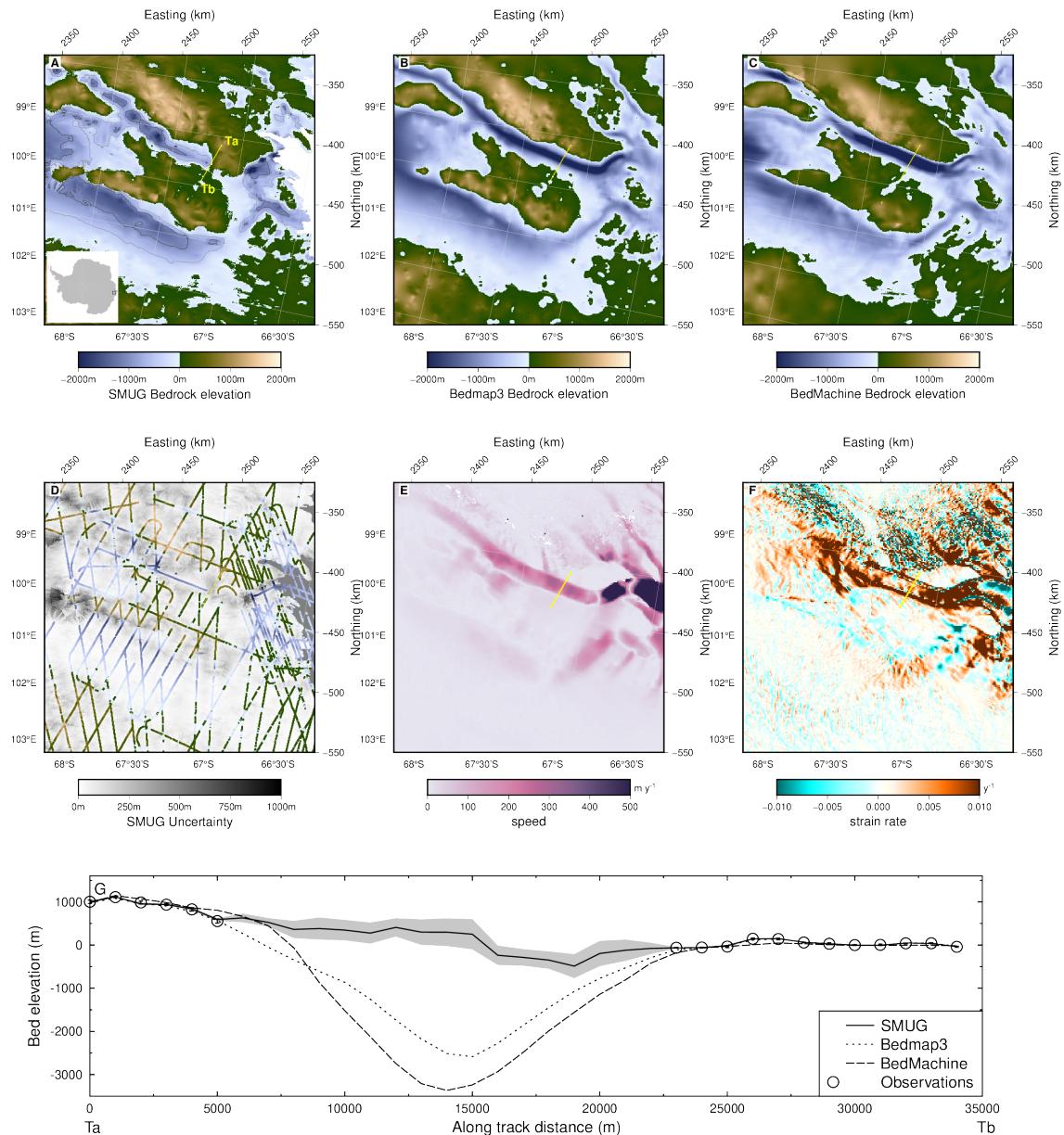


Figure 2: Denman Glacier local region. A) SMUG interpolation, B) Bedmap3 interpolation, C) BedMachine interpolation, D) Bedmap3 observational data, E) MeASUREs surface speed, F) longitudinal strain rate from MeASUREs surface velocity, G) transect Ta-Tb showing SMUG (gray shading shows uncertainty), Bedmap3, BedMachine bed profiles, and the observations.

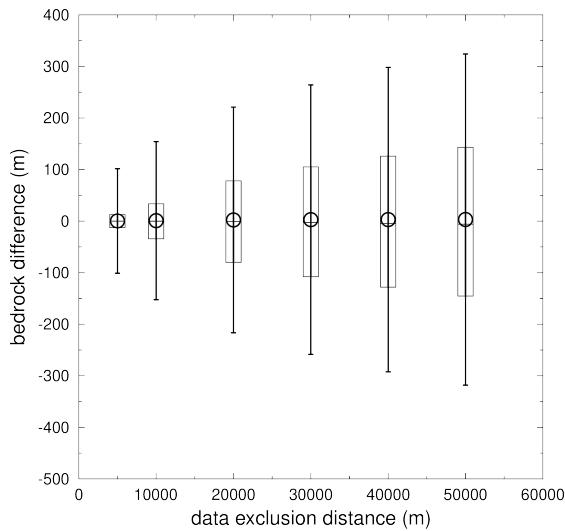


Figure 3: Impact of excluding local data on the interpolation compared to no data exclusion at 1km resolution. Shown are the mean difference (circle), \pm one standard deviation (error bars) and lower quartile, median and upper quartile (boxes).

Stochastic Meshless Uncertainty Gridding

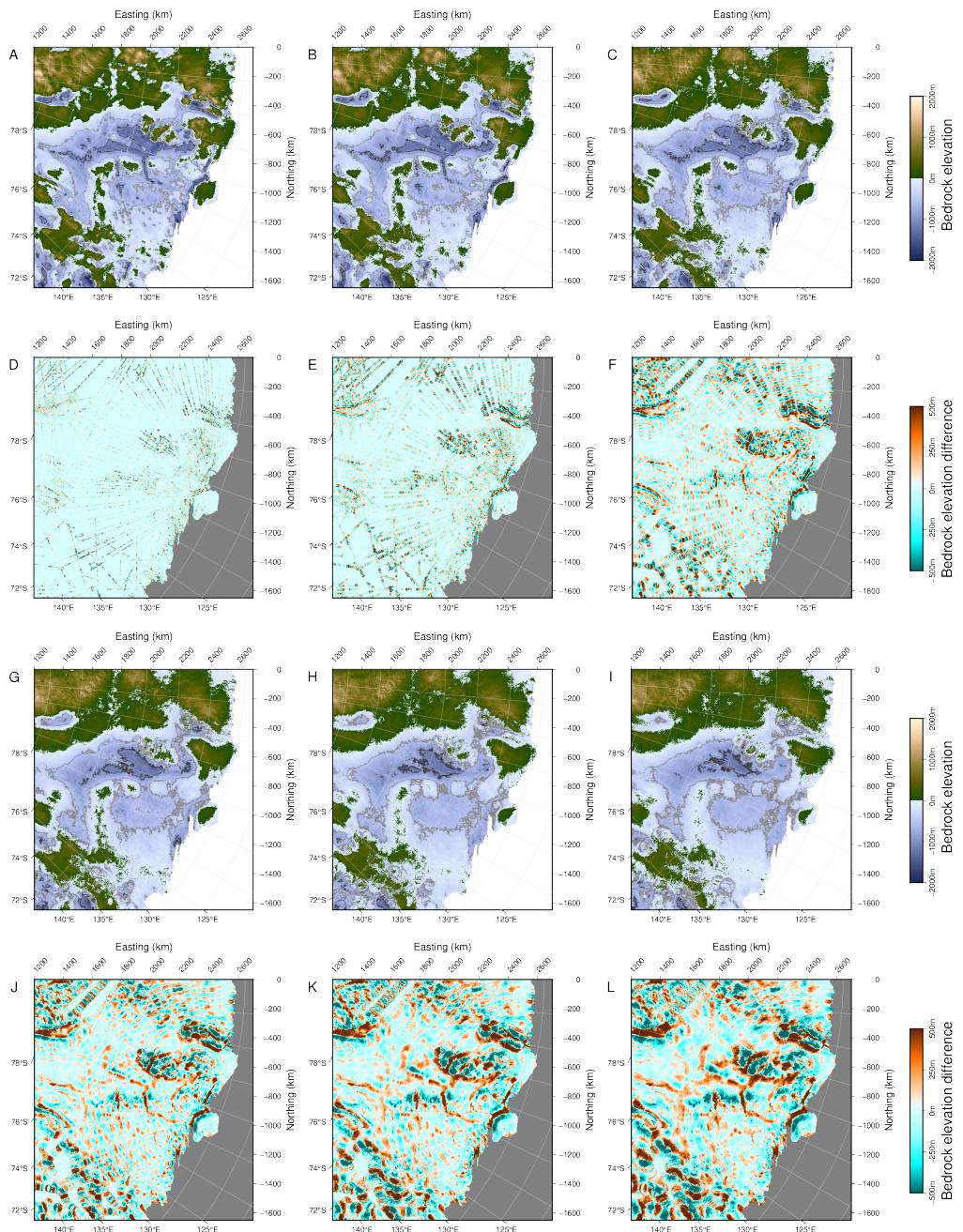


Figure 4: Antarctic bed topography (first and third rows) at 1km horizontal resolution, with observational data excluded, and corresponding difference from interpolation without data exclusion (second and forth rows). Data exclusion radius 5 km (A and D), 10 km (B and E), 20 km (C and F), 30 km, (G and J), 40 km (H and K) and 50 km (I and L).

Stochastic Meshless Uncertainty Gridding

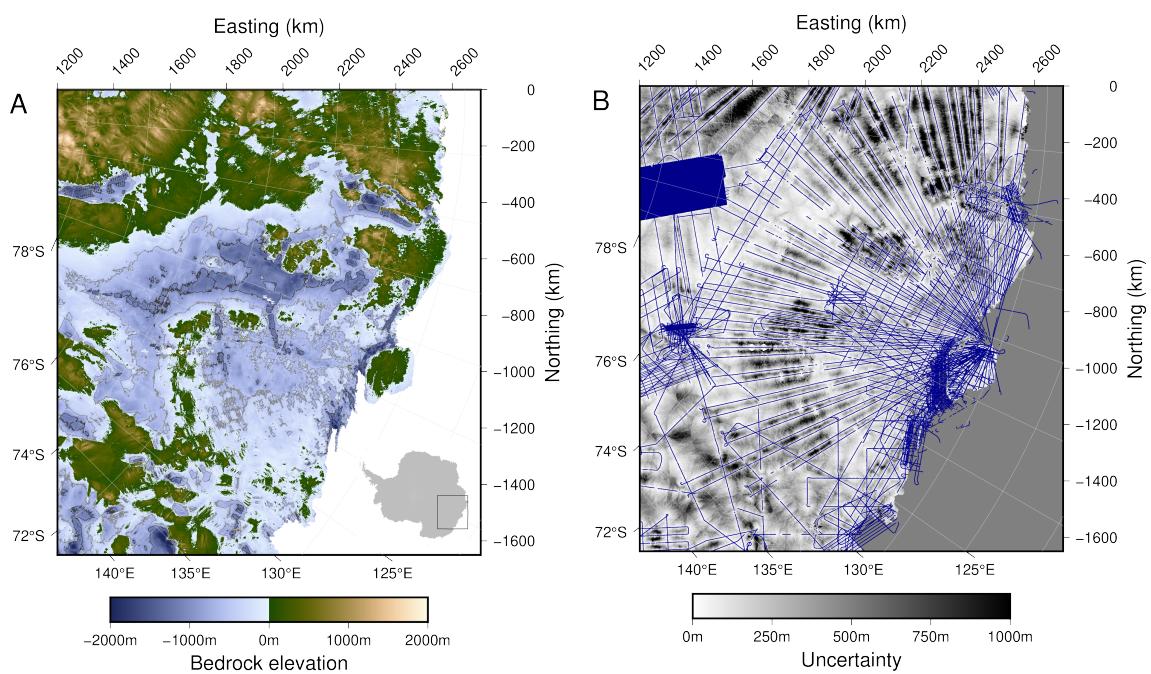


Figure 5: 2nd order Taylor interpolation of ASB at 1 km reconstruction, raw interpolation. A) Bedrock topography, and B) interpolation uncertainty (grey-scale) and observational data locations (blue points).

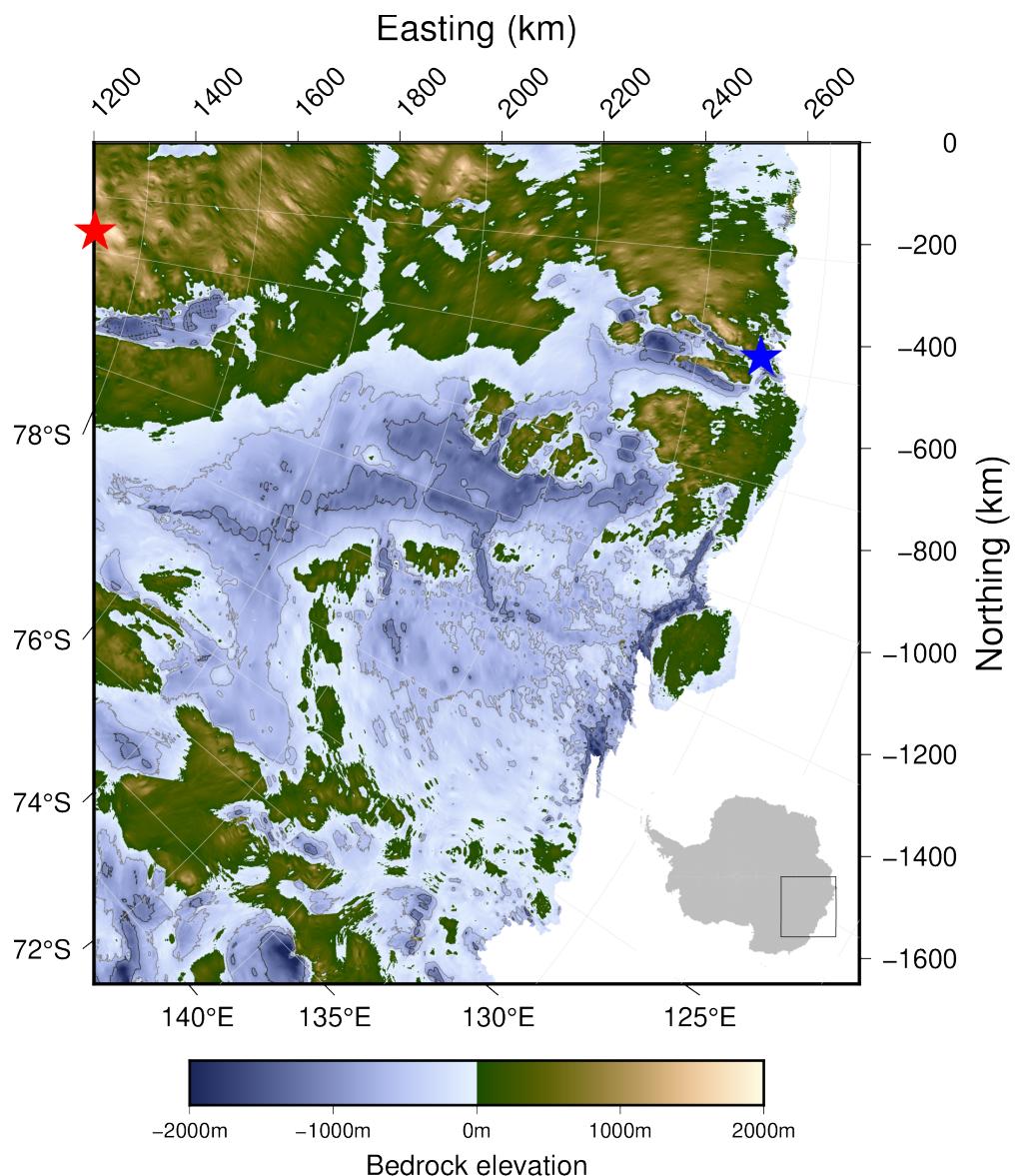


Figure 6: The bed topography dataset after a median filter has been applied to remove impulse noise. The blue star indicates the location of the deepest topography and the red star the highest point in the bedrock.

Stochastic Meshless Uncertainty Gridding

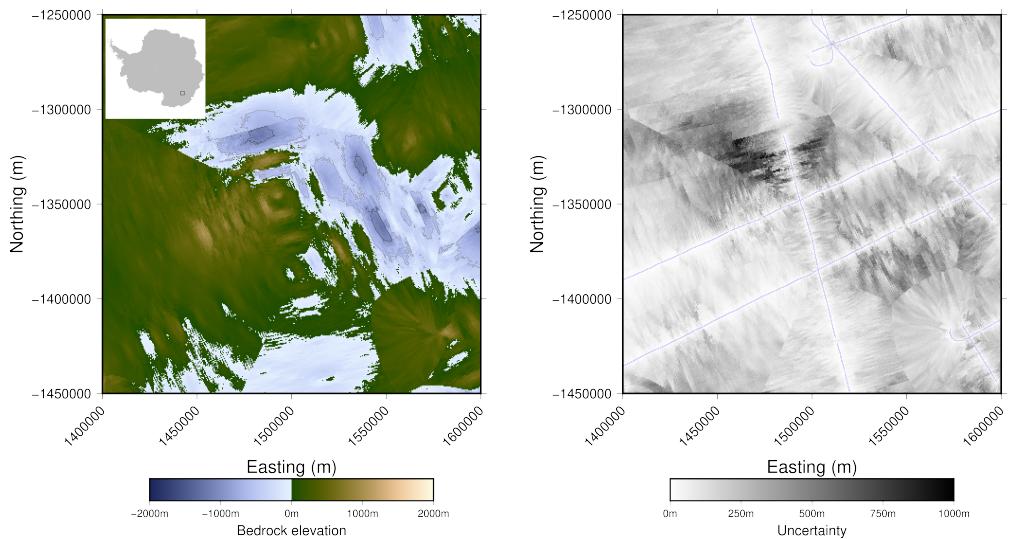


Figure 7: Example of anisotropic effects due to distribution of observations, noting the striping present and discontinuity where there are fewer observations.