UNCERTAINTY ESTIMATION FOR SPARSE DATA GRIDDING ALGORITHMS

Samantha J. Zambo, Paul A. Elmore, A. Louise Perkins, and Brian S. Bourgeois

Naval Research Laboratory, Marine Geosciences Division, Stennis Space Center, MS 39529

ABSTRACT

Fast approximation algorithms that map sparse, irregularly spaced data to a regular grid typically do not include uncertainty estimates. Yet the synthesis of uncertainty with gridded output surfaces is now a required practice. In this paper, we extend the work of Calder and Mayer by including the influence of navigational uncertainty over a sloping seafloor. To do this, we use a circumscribed, nearest neighbor algorithm to determine which data points will contribute to the uncertainty estimate for each gridded output location. Two case studies, one with sparse synthetic data and one with real data around the region of Svalbard, demonstrate the utility of this method. To study the nature and influence of sparse data, the synthetic case study generates a randomized sparse data set whereas the latter case study uses an irregular and sparse data set. The real data case study compares well with previous studies in the Svalbard region and the synthetic case study displays reasonable output values.

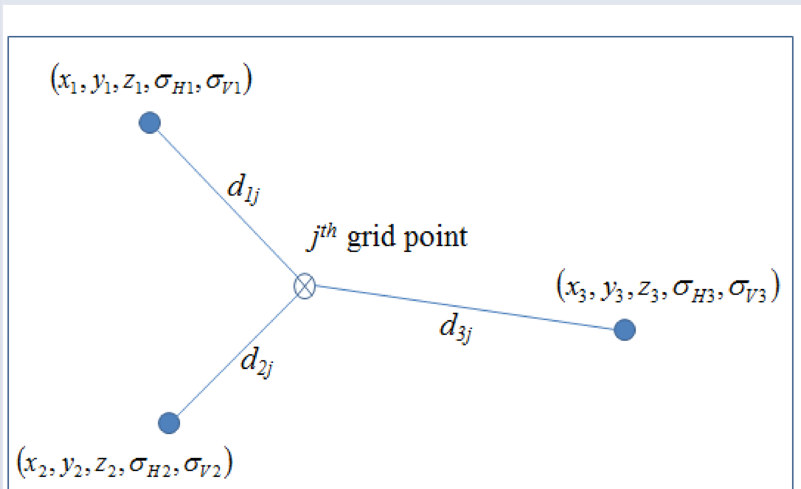
I. PROBLEM

We need a gridded uncertainty estimate for sparse data gridding algorithms that often lack a native uncertainty estimator.

II. SOLUTION: CURVE

(CUBE Uncertainty pRopagated Variance Equation)

1. Estimate Uncertainty : CUBE's Propagated Variance Equation [1]
2. Account for Bottom Slope : Augment with bottom slope term
3. Adapt to Sparse Data : Delaunay TINs



CURVE from (2):

$$\sigma_{ij}^2 = \sigma_{V,i}^2\left(1+\left[\frac{d_{ij}+1.96\sigma_{H,i}}{\Delta_{grid}}\right]^2\right)+\sigma_{H,i}^2\tan^2\theta_j$$

Step 1:

1. Create TIN from input positions
   a) For each j^{th} grid point, find encompassing triangle
   b) Compute d_{ij} from each vertex

Step 2:

1. Compute slope at j from 3rd-order differences of gridded bathymetry
2. Get σ_{ij} from (2)

Step 3:

1. Final σ_j = IDW($\sigma_{1j}, \sigma_{2j}, \sigma_{3j}$)

III. EQUATIONS & COMPONENTS

CUBE's Propagated Variance Estimator [1]:

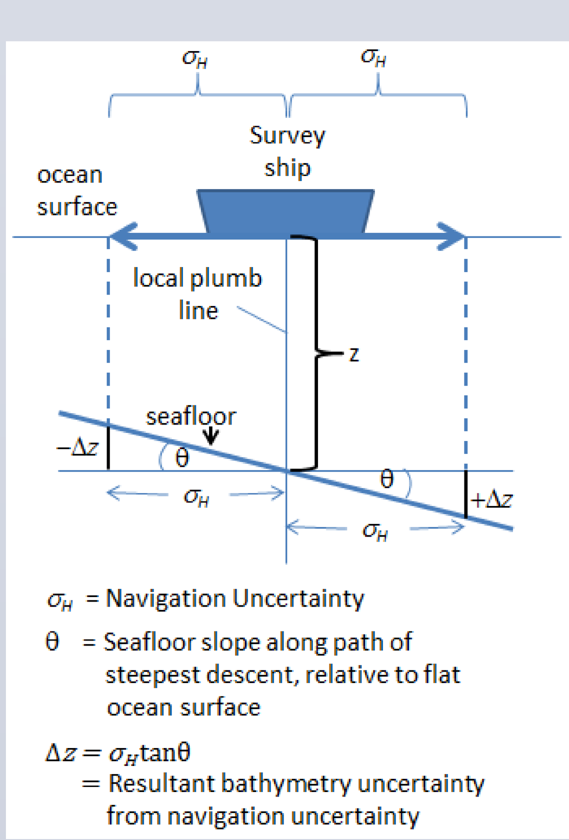
$$\sigma_{ij}^2 = \sigma_{V,i}^2\left(1+\left[\frac{d_{ij}+S_H\sigma_{H,i}}{\Delta grid}\right]^\alpha\right) \quad (1)$$

CURVE:

$$\sigma_{ij}^2 = \sigma_{V,i}^2\left(1+\left[\frac{d_{ij}+S_H\sigma_{H,i}}{\Delta grid}\right]^\alpha + \sigma_{H,i}^2\tan^2\theta_j\right) \quad (2)$$

Bottom Slope:

- The figure to the right shows the relationship between both slope and (2).
- The additional uncertainty term is computed using the assumptions in this figure.
- Θ estimated using third-order finite differences [2] [3].



Inverse Distance Weighting (IDW):

$$\sigma_j^2 = \frac{\sum_{i=1}^{3} d_{ij}^{-1}\sigma_{ij}^2}{\sum_{i=1}^{3} d_{ij}^{-1}} \quad (3)$$

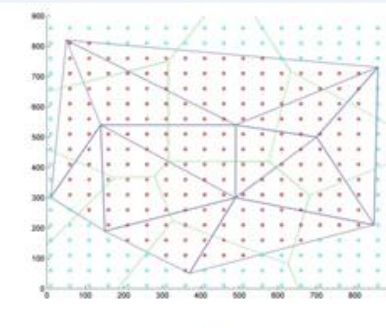
IV. TEST CASE RESULTS

A. Synthetic Case Study

(a-b) displays input data: (a) randomly generated input data with noise from output grid; (b) Delaunay TIN of input data with Voronoi outlines and surface grid points. (c-e) display results: (c) tabulated output gridded data with intermediate computations and total uncertainty estimation 'Uncert' (d) output gridded depth; (e) output gridded uncertainty.

Input X	Input Y	Input Z	Unc_horr	Unc_Vert	grid spacing
850.119	210.8909	554.9	20	3	50
860.1626	730.6991	545.7	20	3	
140.2238	540.2575	513.3	20	3	
370.276	50.7513	550.7	20	3	
490.3404	300.1386	586.1	20	3	
10.4984	300.9593	513.9	20	3	
150.5853	190.1493	528.5	20	3	
50.6551	820.506	521.5	20	3	
490.6797	540.2551	557.3	20	3	
700.9597	500.5472	565.3	20	3	

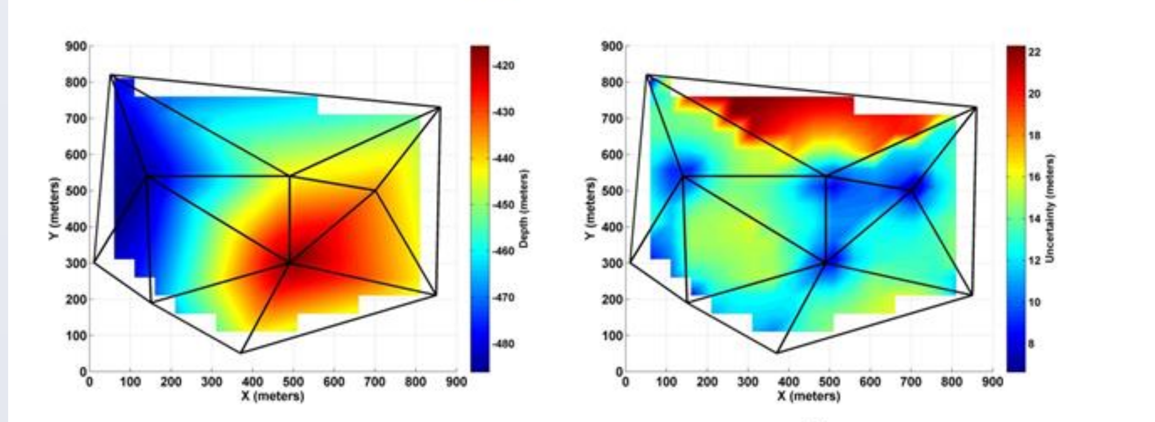
(a)



(b)

Test_X	Test_Y	Test_Z	Slope	TIN_Idx 1	TIN_Idx 2	TIN_Idx 3	d1	d2	d3	var_1	var_2	var_3	Uncert
310	360	551.029	10.34804	3	7	5	250.48	234.03	193.23	285.71	254.65	186.01	15.41
110	360	518.5815	5.310107	7	3	6	174.91	182.85	115.79	149.22	160.59	78.84	11.03
60	510	515.7286	0.989458	6	3	8	214.83	85.77	310.70	207.64	49.40	402.82	12.06
410	210	567.8964	8.334248	7	4	5	263.14	165.03	122.11	306.19	140.83	90.29	12.36
360	560	542.263	6.727287	3	9	8	222.55	133.02	404.96	226.36	98.86	664.68	15.33
810	560	553.4231	5.427318	10	1	2	124.76	351.41	178.08	88.05	509.21	153.87	13.54
510	660	545.951	5.692662	9	2	8	121.82	357.23	487.19	85.38	525.26	939.06	17.72
510	210	570.4036	9.310095	5	4	1	93.58	212.77	340.47	66.19	214.81	487.54	13.11
610	560	557.6711	5.81998	9	5	10	120.94	287.50	108.93	84.67	353.55	73.00	11.15
360	510	544.5274	8.549922	9	3	5	134.74	224.04	250.52	104.24	232.43	281.49	13.60
710	610	554.6837	5.381123	9	10	2	230.16	110.34	192.87	237.83	73.71	175.67	11.85
810	260	557.9862	4.575868	5	1	10	323.39	63.49	264.21	436.07	36.66	302.35	11.65

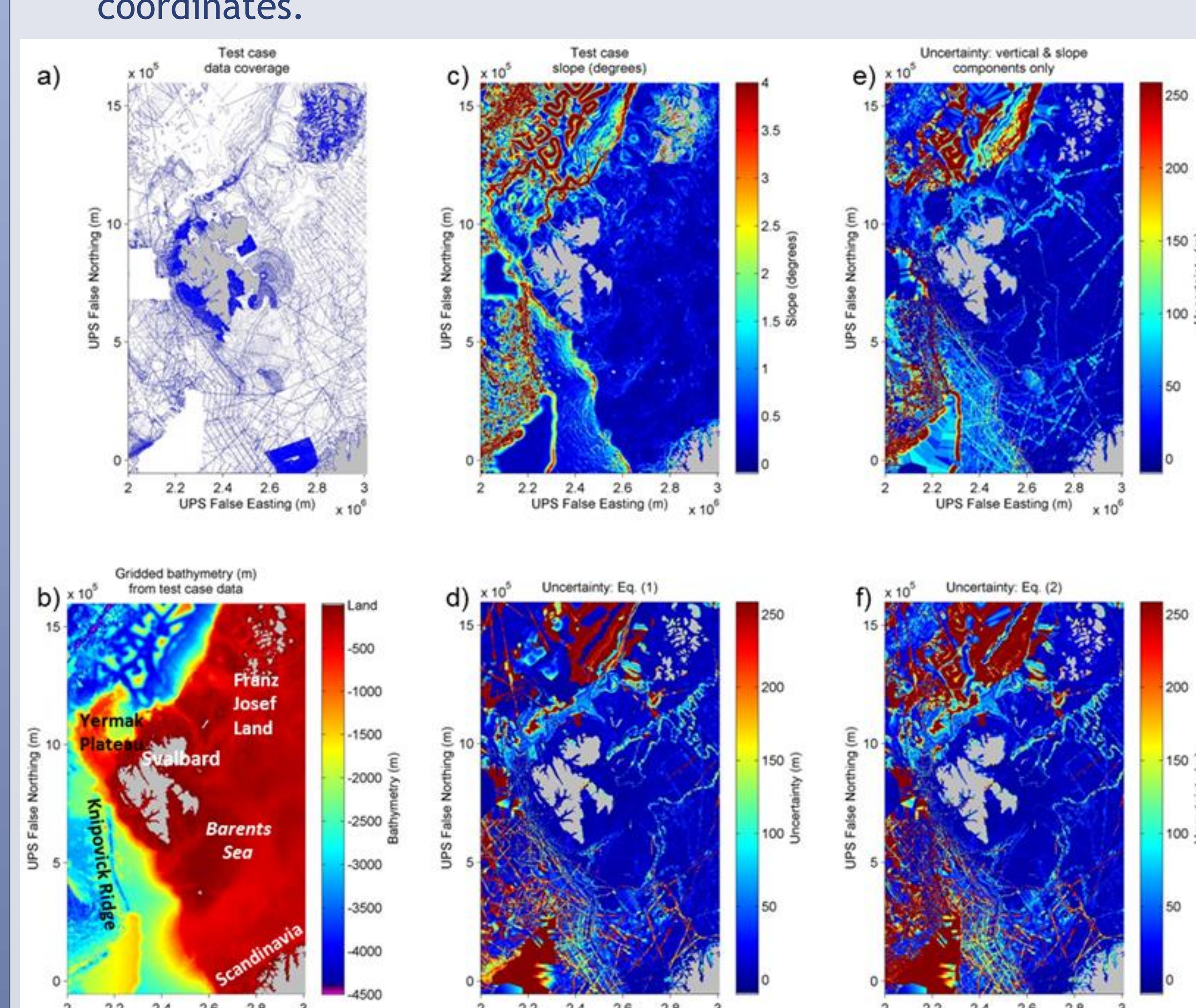
(c)



(d) (e)

B. Svalbard Regional Bathymetry Case Study

(a-c) indicate input data: (a) 2.5km gridded bathymetry of Svalbard region; (b) map of data coverage, where each point is an input data location; (c) gridded slope as calculated from (a). Figs. 6(d-f) are three gridded uncertainty estimates plotted to the same color scale: (d) using (1) and maximum color set at 90th percentile; (e) omitting the middle term in (2); (f) the full estimate from (2). All maps are in Universal Polar Stereographic coordinates.

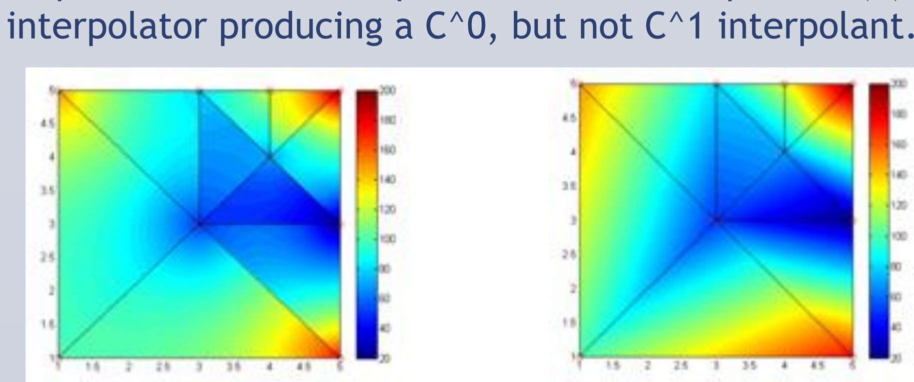


V. FUTURE WORK

- When vertices ≠ nearest neighbors, we will have C^1 discontinuities.
- Oceanographic models utilize bathymetric gridded surfaces as forcing terms which may lead to significant high frequency noise that is slow to dissipate, as found in [5].
- Our ongoing research includes developing methods that extend the number and nature of the contributing vertices.

IDW vs linear interpolation:

(a-b) show discontinuities along triangle edges in Right TINs: (a) an IDW interpolator that fails to produce a C^0 interpolant; (b) a linear interpolator producing a C^0, but not C^1 interpolant.

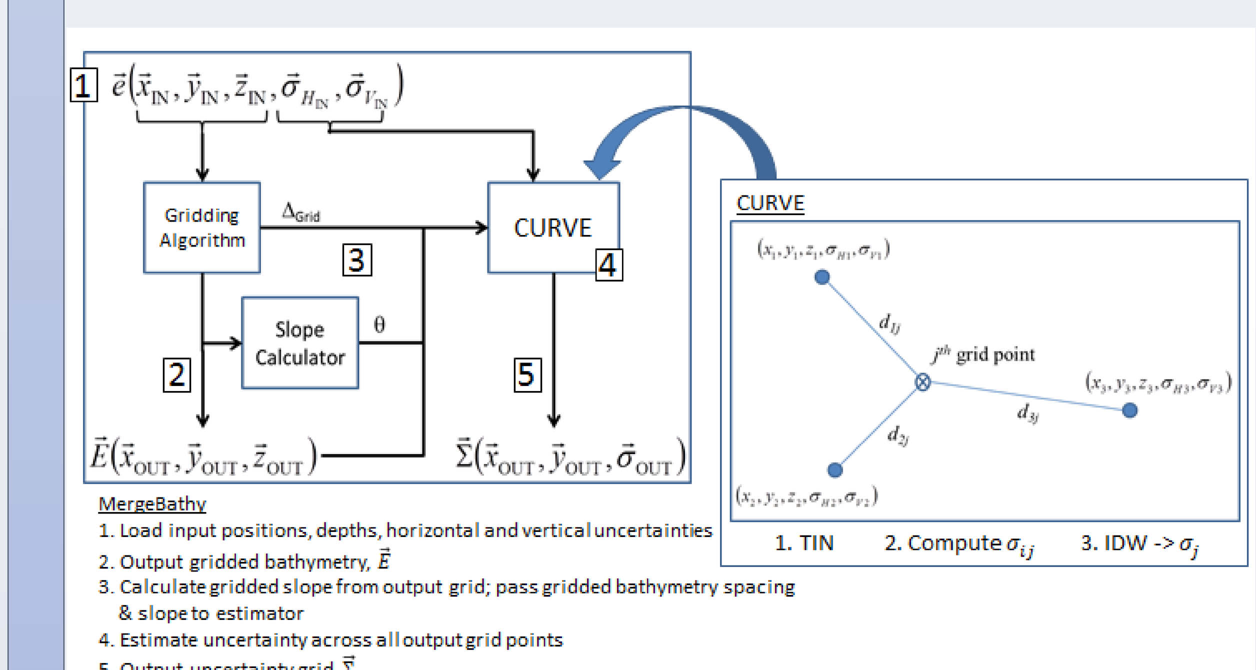


(a) (b)

VI. CONCLUSION

We propose an uncertainty estimator for gridding algorithms that lack an inherent uncertainty estimator. We augment the zeroth-order CUBE uncertainty estimator to handle both sparse data and seafloor slope. First, we append a term for seafloor slope to CUBE's propagated variance equation to form our CURVE algorithm, and then apply it in a way appropriate for sparse data. CURVE uses a Delaunay TIN for selecting control points, propagating their uncertainties, and uses IDW averaging to aggregate contributing uncertainties. The result is a total uncertainty estimate for each gridded surface point. We presented two case studies, the first with synthetic data to demonstrate slope and distance effects on uncertainty, and the second with real data demonstrating CURVE.

MergeBathy's algorithm flowchart for utilizing CURVE to obtain a gridded bathymetric surface with gridded uncertainty from sparse data.



REFERENCES

[1] B. R. Calder and L. A. Mayer. (2003). Automatic processing of high-rate, high-density multibeam echosounder data. *Geochemistry Geophysics Geosystems*. [Online]. *4(6)*. Article. 1048. Available: http://dx.doi.org/10.1029/2002GC000486.

[2] B. K. Horn. (1981). Hill shading and the reflectance map, *Proceedings of the IEEE*. [Online]. *69(1)*. pp. 14-47. Available: http://dx.doi.org/10.1109/proc.1981.11918.

[3] Q. M. Zhou and X. J. Liu, (2004). Error analysis on grid-based slope and aspect algorithms. *Photogrammetric Engineering and Remote Sensing*. [Online]. *70(8)*, pp. 957-962. Available: http://dx.doi.org/10.14358/pers.70.8.957 http://dx.doi.org/10.1016/j.cageo.2003.07.005

[4] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes : The Art of Scientific Computing*, 3rd ed., Cambridge, UK : Cambridge University Press, 2007.

[5] A. L. Perkins. (1995). Regional Gulf Stream initial conditions for the SPEM Primitive Equation Model. *Marine Technology Society Journal*. *29(2)*, pp. 15-23. Available: ISSN: 0025-3324.

ACKNOWLEDGMENT & CONTACT

The authors are very grateful to Dr. Martin Jakobsson (Stockholm U. in Sweden) and Dr. Brian Calder (U. of New Hampshire) who provided us with the data they used for their 2002 paper on the Monte Carlo procedure and technical guidance. We also thank Dr. Calder and Dr. Fred Petry (NRL) for their helpful discussions and feedback on the paper.

Submitted 20 February 2015 to US Hydro 2015. This work is sponsored by the Office of Naval Research through the "Irregular Multiresolution Database Algorithm" Base Program Project at the Naval Research Laboratory. The mention of commercial products or the use of company names does not imply in any way endorsement by the U.S. Navy.

S. J. Zambo is a Doctoral Research Fellow with the Marine Geosciences Division, Naval Research Laboratory, Stennis Space Center, MS 39529 USA (e-mail: samantha.zambo@nrlssc.navy.mil).

P. A. Elmore is with the Marine Geosciences Division, Naval Research Laboratory, Stennis Space Center, MS 39529 USA. (e-mail: paul.elmore@nrlssc.navy.mil).

A. L. Perkins is a Professor with the University of Southern Mississippi and also with the Marine Geosciences Division, Naval Research Laboratory, Stennis Space Center, MS 39529 USA. (e-mail: louise.perkins@usm.edu).

B. S. Bourgeois is with the Marine Geosciences Division, Naval Research Laboratory, Stennis Space Center, MS 39529 USA. (e-mail: brain.bourgeois@nrlssc.navy.mil).