

Modeling of Anticipated Subsidence due to Gas Extraction Using Kriging on Sparse Data Sets

Matthew TAIT and Andrew HUNTER, Canada

Key words: Kriging, Trend Surfaces, Sparse networks, Subsidence monitoring

SUMMARY

This paper investigates different networks of survey monuments to determine whether surfaces fitted using trend analysis and Kriging can replicate a complex predicted deformation surface, and therefore monitor the progress of the subsidence over time. The network configuration was complicated by the sparse availability of land in the subsidence zone, on which to place survey monuments. The study shows that the best network configuration can be determined but that, using this method there is little point in undertaking an annual campaign of measurement.

Modeling of Anticipated Subsidence due to Gas Extraction Using Kriging on Sparse Data Sets

Matthew TAIT and Andrew HUNTER, Canada

1. INTRODUCTION

A feasibility study into the methods of monitoring permafrost deformation due to gas extraction in the Mackenzie Delta of Canada has previously been reported (Tait and Moorman, 2003). In that paper, the initial findings of an investigation into the use of Kriging to represent a complex subsidence bowl were introduced. The main drawback of that study was identified as the lack of reliable variance estimates, for both the chosen observation method and the stability of survey monuments in permafrost. Studies into both have now been completed, which allows a more complete report into the surface fitting method.

Applying models to deformation parameters has long been undertaken in order to discover the underlying physical mechanism of the movements. In common with other reported work (Kenselaar and Quadvlieg, 2001) the best-fit polynomial surface derived in this study is a purely geometric entity that aims to approximate the surface generated by a geomechanical model. Unlike that study, the volume that is available for measurement in this case allows only a sparse network of observations, and not in the principal areas of the geometric shape. The predicted shape itself is more complex due to a lower ratio of depth to radius for the gas reservoir, plus faulting across the study area. Kenselaar and Quadvlieg noted that perhaps the most important aspect in the application of surfaces was the ability to use relative height parameters, precluding the need for a specific monument to remain stable over the length of the monitoring period. In the case of this study, carrier-phase DGPS was proposed as the observation method, a technique that generally gives better relative accuracy than absolute. DGPS was chosen mainly due to the large water-bodies in the study area (up to 2km), across which optical levelling could become rather inaccurate.

Figure 1 shows the anticipated subsidence bowl configuration and total movement contours for the 'most likely' scenario developed by geophysical packages (Tait and Moorman, 2003). The skewed shape derives from faulting in the North-East to South-West direction. The maximum annual subsidence is anticipated to be around 25mm. The axes and centre of the resulting 'ellipse' fall under the waters of the Middle Channel of the Mackenzie River. The area is made up of deltaic sand and gravel with very low relief, with continuous permafrost and massive ice intrusions in the soil to a depth of up to 600m. The permafrost active layer depth is unknown, but is regionally between 0.5 and 2.0m in depth.

Clearly, there is little of the subsidence area that can be measured on land, yet the complex siltation of the channels in the area makes hydrographic measurement meaningless. Methods of areal monitoring using remote sensing are, with the exception of Differential Interferometric SAR (DInSAR), not sufficiently precise to monitor the small subsidence on an annual or biennial basis. DInSAR in this region brings its own unknown factors, but is currently under investigation as part of the larger project goal. Other studies in the project,

such as carrier-phase differential GPS levelling at 70° North, and the stability of survey monumentation in permafrost have been completed (Tait *et al*, 2004a, Tait *et al*, 2004b). These studies have allowed the extension of this study to include the use of Kriging to analyse the subsidence surface and monitor its progress. This paper discusses the fitting of a zero-measurement error surface to the predicted surface shown in Figure 1, followed by the effect on this prediction surface when the variances of the observation method and survey monument stability are added to the estimation.

2. FITTING THE ZERO-ERROR SURFACE

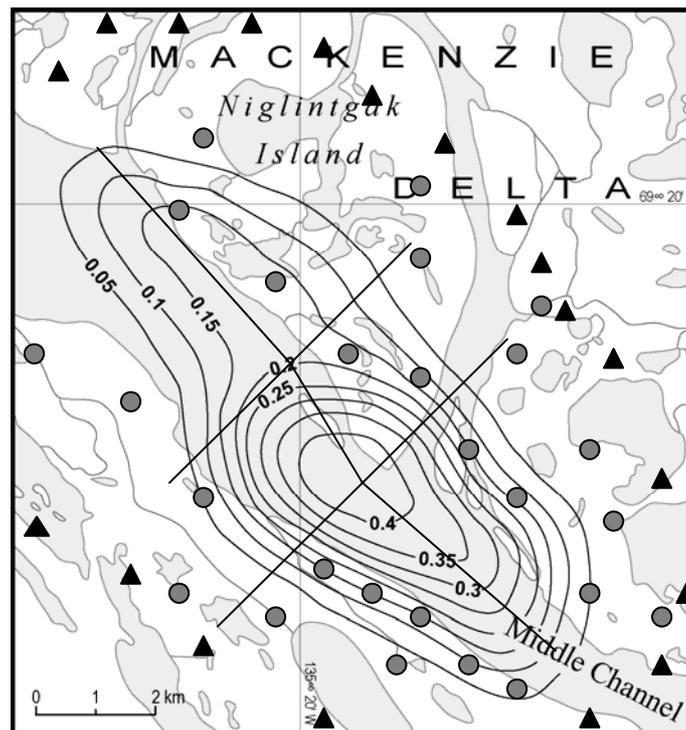


Figure 1: Detail of the Niglintgak Island area showing the predicted subsidence bowl (contours in metres), the spread of monuments in Model A (circles) and the artificial monuments (triangles), and the major axes of the subsidence bowl

This section describes the analysis of four proposed survey networks for the monitoring of subsidence due to the extraction of gas at Niglintgak. The Base Surface against which the four models have been compared was derived from the subsidence model shown by contours in Figure 1. In order to ensure that the same regression model was applied to each survey monument configuration, Trend Surface Analysis was undertaken and all Kriging estimates were derived from the deviations (local autocorrelated component) of the selected model. Residuals for each modelled surface have been determined using the same Trend surface to ensure consistency.

2.1 Survey Monument Network Design

A total of four models were analyzed for this project. The first model consisted of 27 monuments randomly distributed throughout the study area (Figure 1). The three remaining models were based on the same general configuration, but with different numbers of monuments used in each case. In order to be able to model a phenomenon it is extremely important that significant features within the location are captured by the model. With respect to the predicted subsidence, the significant features are low points in each of the depressions and the main axes that define the general shape of the subsidence. These features have been highlighted in Figure 1. Monuments were then placed to coincide as closely as possible with these features. Monuments for Model B were placed close together within the steeper portions of the study area, at approximately 500m intervals radiating out from the centre of each depression (Figure 2).

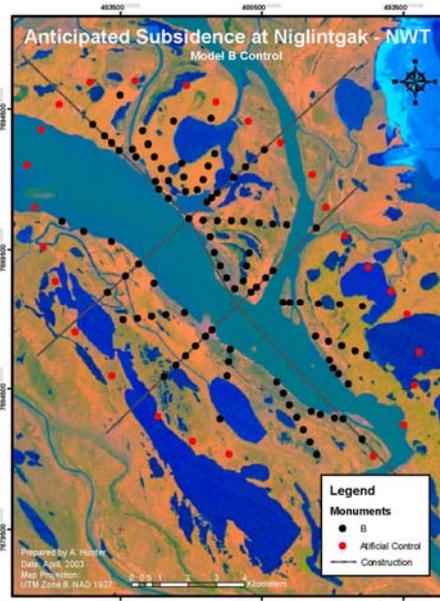
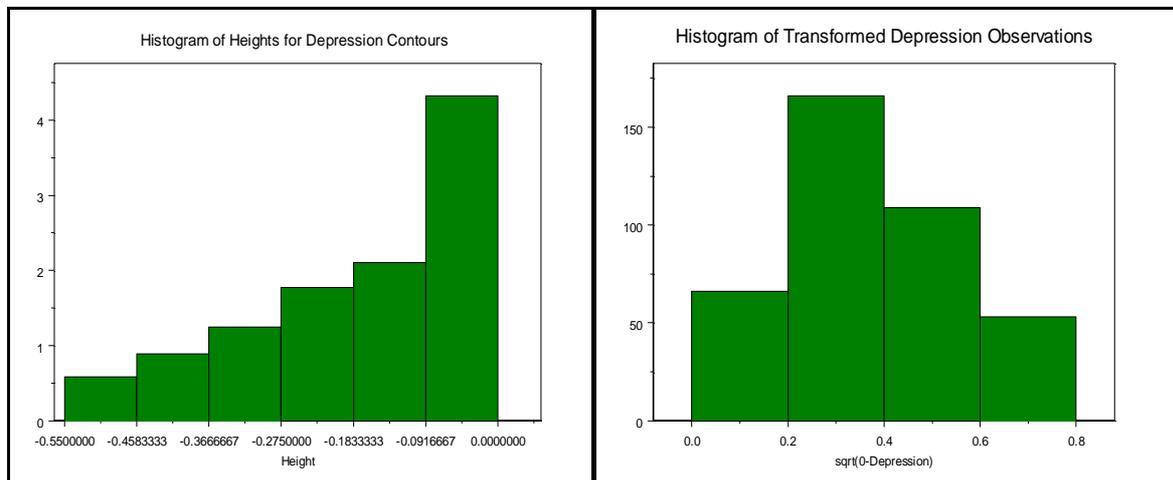


Figure 2: Model B

A total of 100 monuments were used for Model B. Model C was created by removing nearly every second monument from Model B, resulting in a total of 56 monuments being used. A further 13 monuments were removed for Model D. The primary difficulty in the placement of monumentation is that the centre of the larger depression falls within the estuary that runs through the study area. This results in additional control being placed along the edge of the estuary in an attempt to improve the shape of the predicted surface. The heights assigned to each monument used in the Kriging process have been obtained from the estimated Base Surface. In all predicted models, a buffer of artificial monuments was included to limit the shape of the surface being modelled.

2.2 Trend Surface Analysis

An initial review of the contour data indicated that the data was negatively skewed (Skewness = -0.66) as shown in Figure 3a, with a mean of -0.18m and a standard deviation of 0.16m, indicating that it would be necessary to transform the data prior to analysis. The exploratory analysis undertaken also indicated that there was a second order trend in the data and that it was anisotropic. While it is possible to utilize Universal Kriging within ArcGIS to account for regional trend within the dataset, ArcGIS does not allow control over the type of trend surface that is used, aside from a first, second or third order surface, nor does ArcGIS allow you to apply a specific trend surface to a number of different data sets. The outcome of this is that it is not appropriate to compare two surfaces in ArcGIS when there is a trend, as different sample sets will generate different surfaces for the same data. The trend was therefore modelled in S-Plus, initially using a third-degree polynomial followed by a stepwise analysis of the terms. The significance of each model was tested at $\alpha = 0.05$ using the F Statistic. If a model was significant, each individual element of the model was tested for significance using the student-t test.



(a) (b)
Figure 3: Histogram of residuals before (a) and after (b) Transformation

The most appropriate trend model (F: 207.5, k = 6, n-k-1= 387, p < 0.0000) was found to be:

$$\sqrt{(0 - \text{Depression})} = -1.723 + 3.730x + 3.859y - 2.287y^2 - 2.328x^3 - 2.562x^2y - 2.342y^2x$$

The t-statistic for β_1 to β_5 were 26.79, 21.76, -13.27, -20.32, -8.54, -8.21 ($\alpha = 0.05$, n-k-2= 386), with p < 0.0000 in all cases.

The Residuals v. Fitted plot continued to indicate that some structure existed, however this can be partly attributed to the type of data used. Trend surface analysis is based on the assumption that observations are randomly distributed, however in this instance observations are constrained by the location of the contours provided. The Scale-Location plot was less

affected by skewed data and as a result showed considerably less structure than the Residuals v. Fitted plot. The Normal Quantile-Quantile (QQ) plot showed minor deviation from the normality assumption for lower values and is accounted for by the slight negative skew (Skewness: -0.0917) of the transformed observations. Given that the spread of the fitted model for the Residual-Fit Spread plots was greater than the spread of the residuals we can conclude that the model has regression coefficients that are influential to the model. This is also reflected by R^2 , which was 0.7629 for this model. Given that the Normal QQ plot forms a reasonable line; that the residual plots are adequate and the R^2 value has increased by some 18%, the transformation has been considered worthwhile.

2.3 Semivariogram Model Selection

Semivariance is a measure of the degree of spatial dependence between observations, and the selection of an appropriate semivariogram is of primary importance in Kriging. The Semivariogram parameters that are of concern are the Nugget, the Range, the Sill, the Lag Distance and Number of Lags (Davis, 2002, Krajewski *et al*, 2001). Because spatial dependence is a function of “nearness” between objects, a Semivariogram was calculated for each model, since the sampling distance between monuments is irregular in this case, average nearest-neighbour distances were used for lags (Krajewski *et al*, 2001). The number of lags per monument was set as less than one-half the largest distance in the data set (Johnston *et al*, 2001), as such the number varied between 7 and 14 according to the different monument network designs. As a result of the global trend being removed the residuals are reasonably isotropic. The Nugget varied from model to model, as did the partial Sill. Analysis determined that the Rational Quadratic Semivariogram provided the best results in terms of the Standardized RMS being closest to 1. If the Standardized RMS equals one this implies that the model is neither overestimating nor underestimating the error. The model resulted in the following statistics for each of the models (Table 1).

Table 1: Predicted Model Errors using a Rational Quadratic Semivariogram

Prediction Errors	Base	Model A	Model B	Model C	Model D
Mean	-0.004	-0.004	-0.010	-0.007	-0.008
Root Mean Square	0.042	0.105	0.066	0.084	0.080
Average Standard Error	0.039	0.127	0.066	0.099	0.111
Mean Standardized	-0.025	-0.023	-0.055	-0.033	-0.038
RMS Standardized	0.616	0.762	0.613	0.672	0.638
Maximum Standard Error	0.010	0.031	0.016	0.019	0.021

The maximum depth for the Base Model was revised to -0.612m. A plot of the estimated error between the Base Surface and the surfaces generated in the four test networks was generated (See Figure 4 for an example plot). Visually, Model B (not shown) best reflected the overall shape of the anticipated subsidence, as expected from the results of Table 1.

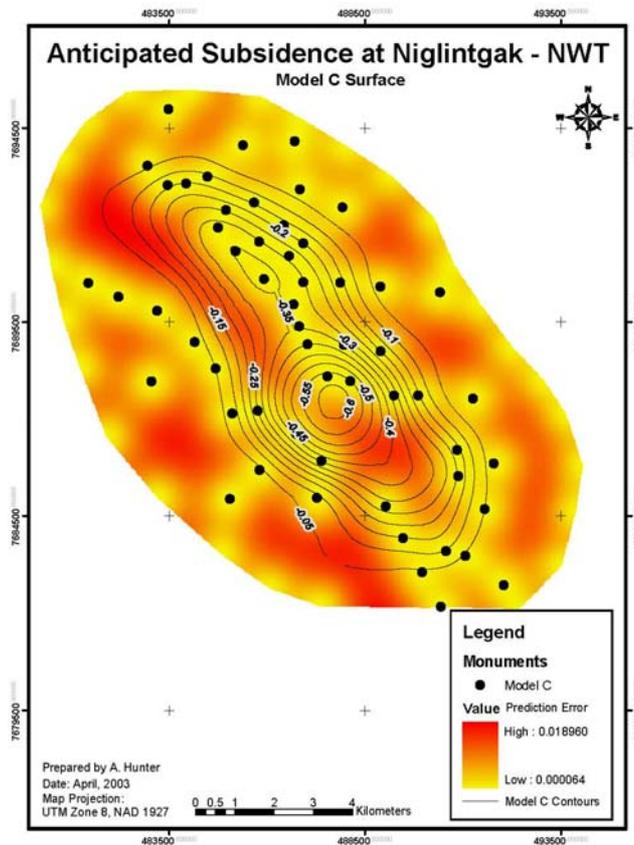


Figure 4: Estimated variance (metres) between Base Surface (contours shown) and Model C

2.4 Comparison of the Different Surfaces

Differences between the modelled surfaces and the Base surface were judged using a 95% Confidence Interval derived from the Kriging error predictions as a means of selecting the most appropriate model. Because error variances are estimated for each model we can calculate a prediction interval for each surface, where the interval:

$$A \equiv (\hat{Z}(s) - 1.96\sigma_k(s), \hat{Z}(s) + 1.96\sigma_k(s))$$

is a nominal 95% prediction interval for $Z(s)$ being the estimated surface, which, under the assumption that $Z(\bullet)$ is stationary and has a normal distribution, satisfies $\Pr\{Z(s) \in A\} = 95\%$.

This analysis has focused on the ability of each survey control network to represent the anticipated subsidence due to the extraction of natural gas. The expectation was that as survey monumentation was removed from the model the ability to predict the anticipated subsidence would decrease, but it was unknown whether the models were statistically different. Therefore, analysis of variance (ANOVA) methods were used to determine whether

or not there was significant difference in the mean differences of the surfaces. At least one of the means was different ($F_{0.05,3,205012} = 2.6050$, $p < 0.0000$); therefore Fisher's Least Significant Difference $LSD = 0.000184$ was used to determine which means were different. Given the model means:

$$\bar{y}_{1\cdot} = 0.00628 \text{ (A)}, \bar{y}_{2\cdot} = 0.00055 \text{ (B)}, \bar{y}_{3\cdot} = -0.00122 \text{ (C)}, \bar{y}_{4\cdot} = -0.00349 \text{ (D)}$$

all paired differences were significant, with model B giving the best result (closest to zero) followed by Models C, D and A.

Table 2: Difference between Models and Base Model at 95% CI (25m cells)

Model	Cells Outside 95% CI	Cells Inside 95% CI	Total Cells	Percent Outside CI	Maximum Difference
A	66,121	138,896	205,017	32.2%	0.190
B	76,781	128,236	205,017	37.4%	0.088
C	76,130	128,887	205,017	37.1%	0.082
D	85,863	119,154	205,017	41.9%	0.068

Table 2 summarizes the differences between each of the models and the base model in terms of area and maximum difference using the 95% CI equation. All models indicate varying degrees of underestimation of subsidence around the low point of the southeast depression. In this area Model A exhibits the greatest deviation at 19cm, while Model D has the least deviation at 6.8cm. Model A also indicated some overestimation of the northwest portion of the northwest depression by 2 to 3cm and under-estimation by approximately 4cm to the northeast of the northern depression. Models B, C and D showed similar differences around the northwest depression; however they are all within 1 or 2cm of the base surface. Model D has the lowest maximum difference when compared to the Base Surface; however, it is assumed that this is partially due to Model D having a larger Average Standard Error (45 mm larger than Model B, and 22mm larger than Model C and therefore a larger Confidence Interval envelope. As mentioned earlier these effects, or differences between surfaces, are due to the inability to place monumentation in appropriate locations due to the estuary that passes through the study area.

Based on this analysis each of the models tested gave significantly different results, and all were significantly different from the Base Surface that the project was attempting to recreate. The best average representation was Model B, followed by Model C. Given that the total area that falls outside the 95% confidence envelope was similar for Models B and C, and that the maximum deviation from this envelope is less for Model C, it was decided that Model C was the more appropriate model, given that there are only 56 monuments required, as opposed to 100 monuments for Model B. Figure 5 shows the predicted differences for Model C using a 95% confidence envelope.

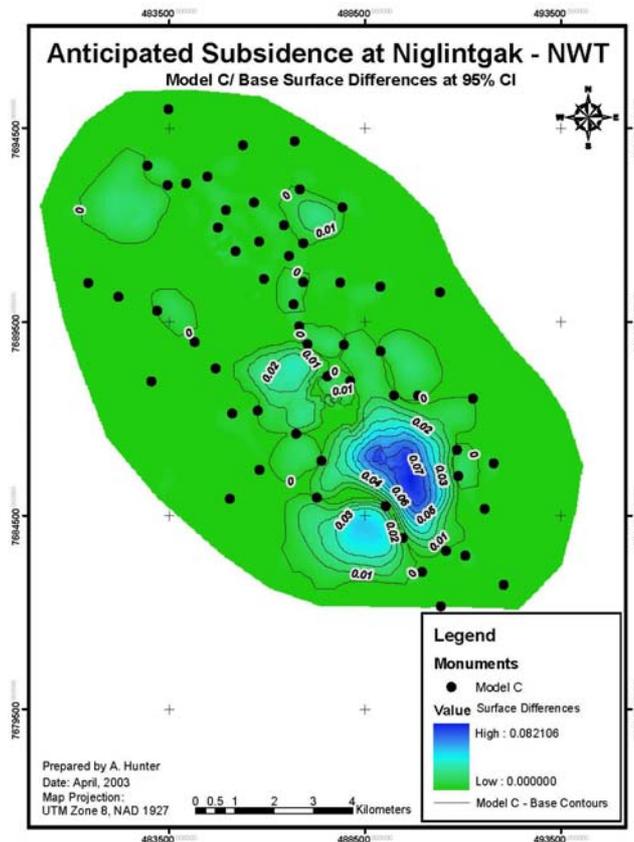


Figure 5: Base - Model C (Rational Quadratic) at 95% CI

3. INCORPORATING VARIANCES IN THE CHOSEN SURFACE

The previous section examined whether a zero-error surface could reproduce the predicted geomechanical surface at different epochs. This section introduces variances for the chosen observation method of DGPS, and the stability of survey monuments in permafrost. As outlined by Cressie *et al* (2003), micro-scale variation (the so called ‘Nugget effect’) causes a discontinuity at the origin which mathematically should not exist. Therefore, if continuity of a phenomenon is expected at the microscale, then the nugget should represent measurement error. But because we are limited by the data used in the calculation of the variogram we are unable to determine if microscale variation is continuous or not. Typically, it is assumed that microscale variation is not continuous, and that the Nugget represents both a microscale process and a measurement error. Therefore, measurement errors are transferred to errors in the specification of the spatial dependence structure used by the Kriging process. If we know what the measurement error is then, following Cressie *et al* (2003), it is possible to remove the effect of measurement error from Kriging prediction errors, thereby improving the quality of the prediction.

Based on fieldwork undertaken during the summer of 2004 (see Tait *et al*, 2004a) the standard deviation for differential heights determined by DGPS in the region was found to be

0.008m. The stability of permafrost monumentation was also investigated (Tait *et al*, 2004b) with the most stable configuration giving around 5mm standard deviation over 17 years. Model C was revisited so that measurement error could be incorporated into the estimation of the surface. It is important to note that it was necessary to use a Gaussian Semivariogram model in this instance as the Nugget calculated from the Rational Quadratic Semivariogram was less than the measurement error variance (0.000032 versus 0.000064). This resulted in an increase in Nugget effect to 0.006629, with a Partial Sill of 0.044194 and a Range of 8151. Lag distance and number of lags were maintained at 920m and 9. This result would indicate that the Rational Quadratic Semivariogram will tend to underestimate the micro-scale variation in this instance. In addition, changing the model of spatial dependence makes it difficult to compare the results from this section with the results described earlier, as the Gaussian model tends to be used to represent samples that have strong regional continuity, whereas the Rational Quadratic model tends to be used to represent samples that have a strong short-range continuity, i.e. spatial dependence reduces more rapidly as the Lag distance increases. The table below highlights the differences between the predicted error surfaces with and without measurement error.

Table 3: Prediction Errors for Model C with and without Measurement Error Specified

	Without ME Specified	With ME Specified
Minimum	0.007	0.006
Mean	0.008	0.007
Maximum	0.010	0.008
Standard Deviation	0.0004	0.0004

These results indicate a small improvement in the prediction error by accounting for approximately 1% of the Nugget effect, however it is not significant ($Z = 0.25$, $p = 0.41$). In effect, for this analysis, the scarcity of monuments overpowers the expected measurement error. This knowledge provides the opportunity to estimate a level of measurement error that will maintain accuracy of prediction and at the same time minimize effort required in the field in order to obtain the data. If measurement error is increased, i.e., we relax the measurement specifications, we do not start predicting a significant difference until the standard deviation of measurement error is in the order of 0.03m.

As is evident from comparison of Figures 5 and 6, use of the Gaussian Semivariogram results in a much smoother surface than does the Rational Quadratic Semivariogram. In addition the Gaussian model fails to represent the major depression in the surface adequately. It is presumed that this is due to the effect of a larger Range value. Further analysis indicates that the major subsidence is some 0.25m above the Base surface. This can be improved by reducing the Range used in the Kriging predictions. However, this contradicts the notion of relatedness implied by a semivariogram, in that the Range is determined endogenously from the data, and therefore, has some relevance to the process being studied, rather than imposing some arbitrary Range on the model. Lastly the Gaussian model predicts a larger average estimation error than does the Rational Quadratic. This is evident in the Sill obtained from the semivariograms (0.032 for the Rational Quadratic and 0.044 for the Gaussian).

Table 4: Difference between Model G (Gaussian) and Base Model at 95% CI (25m cells)

Model	Cells Outside 95% CI	Cells Inside 95% CI	Total Cells	Percent Outside CI	Maximum Difference
C	7,458	197,559	205,017	3.6%	0.250

Table 4 indicates that the Gaussian Model has predicted the surface reasonably well; however this can be largely attributed to the average error envelope being 0.04m wider for the Gaussian Model than for the Rational Quadratic model. Figure 6 describes the differences between the Base model and Model C (Gaussian).

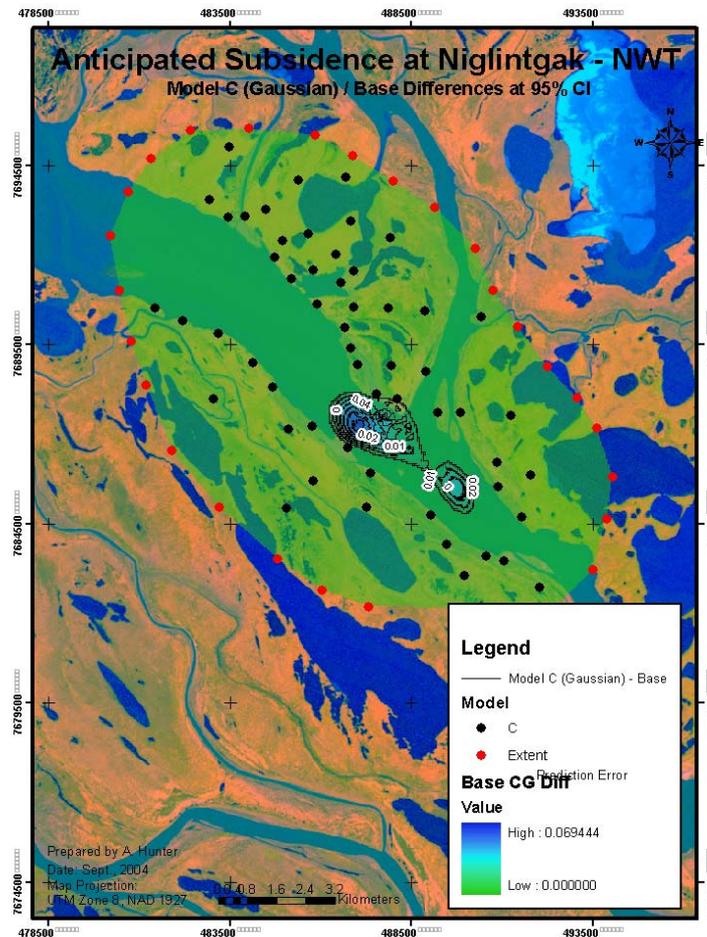


Figure 6: Base - Model C (Gaussian) at 95% CI

4. CONCLUSIONS

It is evident from this work that it is difficult to predict a surface when it is not possible to obtain measurements at significant points. However, the modeling technique provides an unbiased approach for the selection of an appropriate model, and by including measurement error we have the ability to determine the sensitivity of the prediction given a particular network design. One interesting observation, from the introduction of reliable variances into the chosen network model, was that the level of these variances did not alter the prediction surface significantly due to the large distances between monuments having greater effect than the measurement variance. The statistical interpretation of this observation is that a high-precision measurement method is therefore not required for deformation monitoring in this case. The corollary to this, from a geodetic viewpoint, is that the DGPS method will not discover deformation using this approach for, perhaps, three years. Increasing the number of monuments would help this problem. However, the chosen model incorporated over 50 monuments, which is already a substantial number. The real problem remains that data may not be collected in the most appropriate areas for surface modelling.

REFERENCES

- Bailey T. C. and Gatrell A. C. 1995. *Interactive Spatial Data Analysis*, Addison Wesley Longman Ltd, UK
- Burrough P. A. and McDonell R. A. 1998. *Principles of Geographic Information Systems*, Oxford University Press, Oxford UK
- Cressie N. A. C. and Kornal J. 2003. *Spatial Statistics in the Presence of Location Error with an Application to Remote Sensing of the Environment*, Statistical Science, Vol. 18(4), pg. 436 to 456
- Davis J. C. 2002. *Statistics and Data Analysis in Geology*, John Wiley & Sons, New York USA
- Johnston K., Ver Hoef J. M., Krivoruchko K., and Lucas N. 2001. *Using ArcGIS Geostatistical Analyst*, ESRI.
- Kenselaar, F. And Quadvlieg, R. 2001. *Trend signal modelling of land subsidence*. Proceedings of the 10th International Symposium on Deformation Measurements, Orange, California, USA, 19-22 March, 2001.
- Krajewski S., and Gibbs B. 2001. *Understanding Contouring: A Practical Guide to Spatial Estimation Using a Computer and Variogram Interpretation: With Contour Map*, Gibbs Associates.
- Tait M. and Moorman B. 2003. *A feasibility study into monitoring deformation in the Niglintgak region of the Mackenzie delta*. In: Proceedings of the 11th FIG Symposium on Deformation Measurement, 25-29 May 2003, Santorini, Greece, pp.319-326
- Tait M., Moorman B., and Sheng, L. 2004b. *The long-term stability of survey monumentation in permafrost*. Journal of Engineering Geology special volume on 'Geodetic Techniques in Geological Deformation Studies.' (in print)
- Tait M., Sheng L., and Cannon M.E. 2004a. *The feasibility of replacing precise levelling with GPS for permafrost deformation monitoring*. In: proceedings of INGENO 2004 and FIG Regional Central and Eastern European Conference on Engineering Surveying, November 11-13, Bratislava, Slovakia.

CONTACTS

Dr Matthew Tait
University of Calgary
2500 University Drive NW
Calgary
CANADA
Tel. + 1 403 210-9494
Fax + 1 403 284-1980
Email: tait@geomatics.ucalgary.ca