

# **Efficient Cleansing of Radar Rainfall Images**

**Stephen Wesson & Geoffrey Pegram**

Civil Engineering, University of Natal, Durban, 4041, wessons@nu.ac.za & pegram@nu.ac.za

## **ABSTRACT**

Flood Forecasters, and organisations that are affected by rainfall (e.g. Farmers, Disaster Managers, Transport Operators etc.), make increased use of weather radar for estimating and forecasting precipitation. However there are problems associated with the radar rainfall product due to ground clutter and anomalous propagation to just name two. This paper focuses on techniques for correcting the radar rainfall images in a 2-D plane or in a 3-D stack of planes and infilling missing data. The following aspects are highlighted:

- Kriging is considered to be the optimal method for the spatial interpolation of data. The two Kriging techniques, Simple and Ordinary Kriging, are described and how they are implemented. The advantages and disadvantages of the Kriging methods will also be reviewed. The “screening effect” that is associated with the Kriging weightings will be described and how one can take advantage of the distribution of weightings to reduce computational time and increase efficiency.
- The coefficient matrix, which is used to determine the weighting values in Kriging, can be highly ill conditioned depending on both its size and the chosen parameterisation. The method of Singular Value Decomposition can be utilized to resolve this problem in conjunction with the elimination of the near zero singular values. Making economic by reducing the rank of the coefficient matrix by elimination of these small singular values leads to a decrease in computation time and improved efficiency with little or no significant loss in accuracy.

## 1. INTRODUCTION

Raingauge data have been traditionally used for the recording of rainfall over catchment areas and are generally accepted as the true reflection of what the rainfall values are at ground level. There are however limitations to raingauge data, they firstly only provide a point measurement and secondly even with a dense network of raingauges it is difficult to extrapolate and interpolate the rainfall values in any significant detail and accuracy because rainfall accumulation in short times over a surface are not smooth.

Radar rainfall data does address some of the limitations that raingauges exhibit. Rainfall can vary greatly in space and time. Radar rainfall images show the rainfall in far greater spatial detail and complexity than is possible with raingauges. By being able to observe and record the rainfall in far greater detail one is able to predict severe weather patterns in a far more timely and efficient manner. This type of information is invaluable to organisations such as Agriculturalists, Disaster Managers, Catchment Authorities and Transport Operators who depend on accurate and timely information on rainfall. Radar rainfall images provide the spatial information that enables a greater understanding and predictability of rainfall development and movement to be obtained.

There are however problems and errors associated with the radar rainfall images these can include ground clutter, beam blocking, bright band, anomalous propagation and attenuation to name a few. Problems, for example ground clutter (non-meteorological echoes that influence radar data quality), can be caused by the radar beam colliding with the earth's surface. This results in portions of the radar-rainfall images being contaminated with non-rainfall information, these areas can be of significant size and intensity. This is illustrated in Figure 1.1 where the ground clutter is marked in black. In order to obtain the best rainfall estimate possible, the missing rainfall data needs to be estimated by interpolating (or extrapolating) the surrounding available rainfall data.

In order for the radar rainfall data to be of use to the relevant organisations the missing rainfall data needs to be estimated in the most effective and efficient manner as possible and calculated in real time. The proposed method for infilling the missing data is Kriging, which is considered the optimal technique for the spatial interpolation of data. There are however advantages and disadvantages associated with this technique that will be reviewed along with how this technique is implemented.

In order for the Kriging technique to be used as a real time application for infilling the missing or contaminated data, it needs to be as computationally efficient as possible. For example the pattern of weightings around an unknown point can be used with advantage to decrease computation time and improve the efficiency of the infilling technique. Other techniques also need to be found in order to improve the speed of this process.

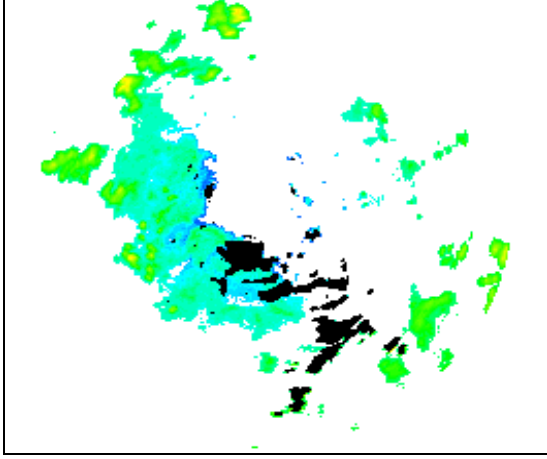


Figure 1.1: Radar rainfall image with ground clutter marked in black

## 2. PROPOSED DATA INFILLING TECHNIQUE

Kriging can be considered to be the optimal technique for interpolating normally distributed data. In Kriging the data set and its structure determine the basis function. In other interpolating techniques such as Multiquadrics the basis function is chosen subjectively for the sake of computational convenience. In this way Kriging takes into account the actual nature and structure of the data set so as to provide the optimal estimate [Pegram, 2001]. The two Kriging techniques that will now be reviewed are Simple and Ordinary Kriging.

In the Simple Kriging technique the mean field is assumed known and estimated from the data set. The equation for Simple Kriging is given below.

$$z(s_0) = M(s_0) + \lambda^T(s_0)[z - M]$$

Where  $z(s_0)$  is the estimated data value,  $M(s_0)$  the known mean value and  $\lambda^T(s_0)$  the vector of weighting values.  $z$  is the vector of known data values and  $M$  is the vector of associated means. The computation is done for each target location  $s_0$ . The weighting values are calculated by the solution of the following matrix equation, when  $g(s_0)$  is a vector of covariances between the target location  $s_0$  and the known data  $z$  and  $G$  is a matrix of covariances between the known  $z$  values.

$$\lambda^T(s_0) = g^T(s_0) \cdot G^{-1}$$

In the Ordinary Kriging equations the mean is assumed unknown. The Ordinary Kriging equation is given below, where again  $z(s_0)$  is the estimated data value at  $s_0$ .

$$z(s_0) = \lambda^T(s_0) \cdot z$$

The weightings are determined from the following equation:

$$\begin{bmatrix} G & u \\ u^T & 0 \end{bmatrix} \cdot \begin{bmatrix} \lambda(s_0) \\ \mu(s_0) \end{bmatrix} = \begin{bmatrix} g(s_0) \\ 1 \end{bmatrix}$$

where  $u$  is a unit vector of ones and  $\mu$  a Lagrange multiplier. In this instance the weightings are constrained to sum to unity whereas in Simple Kriging they are not.

Kriging can be performed using either a semi-variogram or a covariance function in a stationary field, if the latter then both sides of the equation to derive the weights can be divided by the variance in which case  $g(s)$  becomes the correlation function. We choose the semi-variogram as a more stable computational procedure so:

$$g(s) = 1 - \exp\left[-(s/L)^\alpha\right]$$

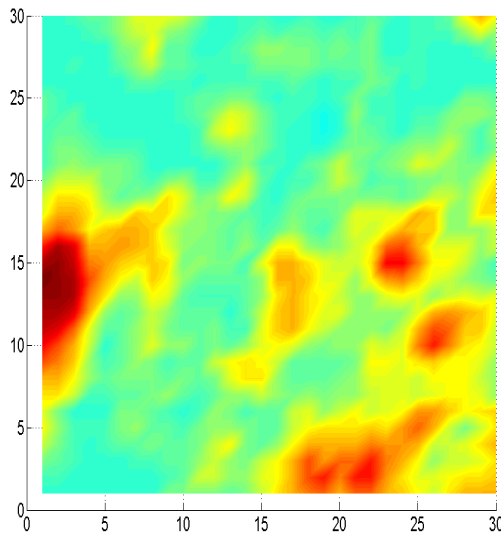
where  $s$  is the distance between known data points,  $L$  is the correlation length, and  $\alpha$  is the exponent value lying in the range (0, 2).

The advantage of the Kriging technique is that it can be easily generalised from a one-dimensional to a two or three-dimensional data set. On extrapolation outside of the data set, the estimated values converge to the mean value of the field whereas other techniques do not exhibit this behaviour. The Kriging Variance also provides the accuracy of the estimated data. The Kriging Variance is calculated by the following equation when a semi-variogram type model is used [Cressie, 1993]:

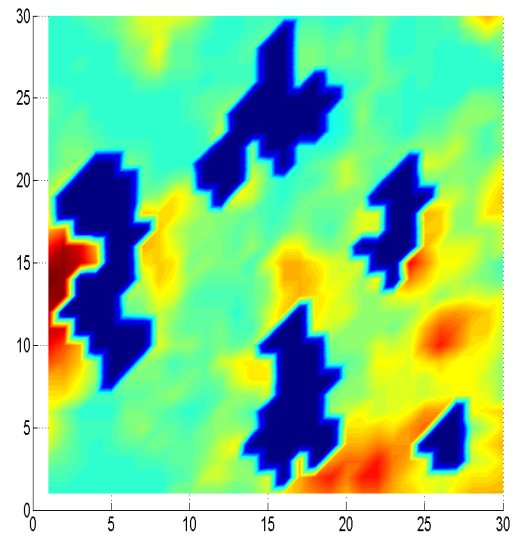
$$\sigma_k^2(s_0) = \lambda^T(s_0) \cdot g(s_0) + \mu$$

The disadvantage of Kriging is that it relies on matrix methods to perform its computations. On large sets of data this becomes burdensome and time consuming. For the Kriging technique to be used in a real time applications it is not feasible to work with matrices above a certain size.

An example of the Ordinary Kriging technique is illustrated on sample rainfall data in Figure 2.1. Segments of the rainfield were removed to reproduce ground clutter effects, indicated by the dark blue segments, and the missing rainfall data estimated using Ordinary Kriging. A semi-variogram model was used with a Correlation Length ( $L$ ) of 11.5 kilometres and an  $\alpha = 1$ . Figure 2.2 indicates the Kriging Variance associated with the estimated data.



Sample 30km by 30km Rainfield



Segments Removed from Rainfield

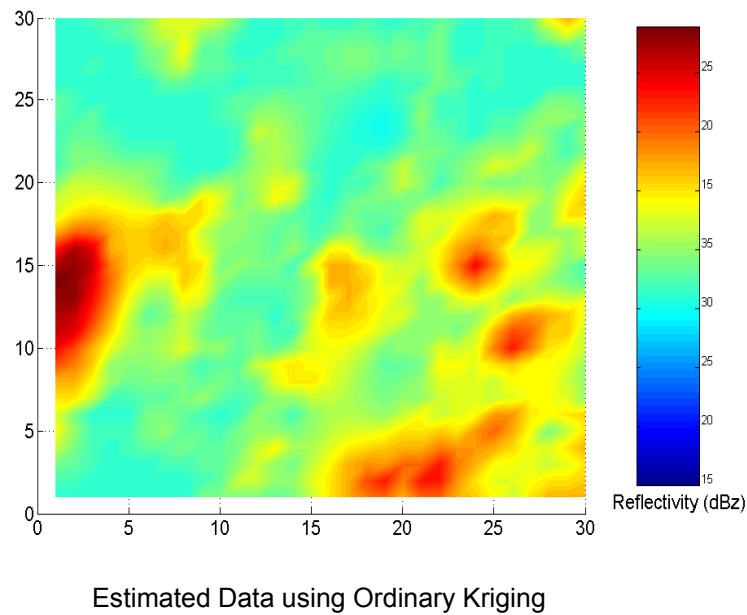


Figure 2.1: Ordinary Kriging used to estimate missing rainfall data

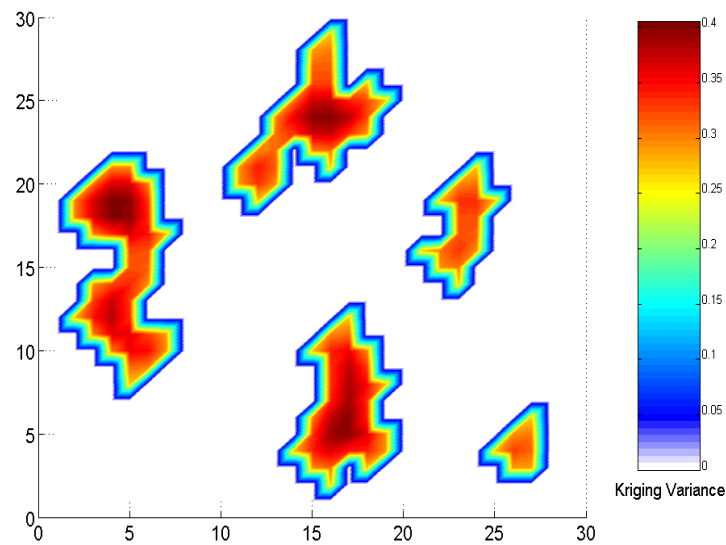


Figure 2.2: Kriging Variance associated with estimated data

In order for the Kriging technique to be used in a real time application it is not feasible to work with matrices beyond a certain size due to limits on computation time and efficiency. However in the Kriging technique applied to dense data sets it is not necessary to use the entire data set in order to estimate the missing information due to the “screening effect”. This occurs when the significant Kriging weights are concentrated around the unknown data points that need to be estimated [Chiles & Delfiner, 1999]. In order to obtain an accurate estimate of a missing data point only the data with weightings of significant values need to be used in the Kriging process.

By only considering these points as oppose to the entire set the computational speed and efficiency is greatly improved. Using the entire data set would be computationally burdensome and the relatively small gain in accuracy would be far outweighed by the time and effort needed

in Kriging from the entire set of non-contaminated values. Only the five to ten nearest neighbours from a contaminated point typically need to be used [Seed & Pegram, 2001]. The points outside the five to ten nearest neighbours have weightings which are significantly close to zero and will have little or no influence on the data point which is being estimated.

Figure 2.3 shows the sum of the Ordinary Kriged weighting values surrounding four unknown data points (indicated by the grey coloured cells). As indicated in Figure 2.3 the significant Kriging weightings are concentrated around the four unknown data points. When a range of greater than two cells from an unknown data point is exceeded the weighting value is approximately zero, these data points can then be excluded in the Kriging process, making for a considerable reduction in the size of the G – matrix and an improvement in efficiency.

0	0	0	0	0	0	0	0	0
0	-0.01	-0.02	-0.04	-0.03	-0.01	0	0	0
0	-0.03	0.04	0.33	-0.01	-0.09	-0.03	-0.01	0
0	-0.06	0.47	•	1.07	0.41	0.04	-0.02	0
0	-0.07	0.52	•	•	•	0.32	-0.04	0
0	-0.03	0.04	0.53	0.64	0.44	0.04	-0.02	0
0	-0.01	-0.03	-0.08	-0.09	-0.07	-0.03	-0.01	0
0	0	0	0	0	0	0	0	0

Figure 2.3: Sum of Kriging Weighting values concentrated around four unknown data points

### 3. SINGULAR VALUE DECOMPOSITION TO IMPROVE EFFICIENCY

In order to determine the weighting values used in the Ordinary Kriging Process the following set of equations needs to be solved.

$$\begin{bmatrix} G & u \\ u^T & 0 \end{bmatrix} \cdot \begin{bmatrix} \lambda(s_0) \\ \mu(s_0) \end{bmatrix} = \begin{bmatrix} g(s_0) \\ 1 \end{bmatrix}$$

However the coefficient matrix can be highly ill-conditioned depending on both its size and the parameters used in the semi-variogram model, it is especially influenced by the  $\alpha$  value that is selected. As indicated in Figure 3.1 as the  $\alpha$  value of the semi-variogram model approaches the value two (i.e. tends from exponential to Gaussian) the matrix becomes increasingly ill-conditioned. Once the coefficient matrix is past a certain size and an alpha value of two is used the solution to the weightings in the above equation are meaningless.

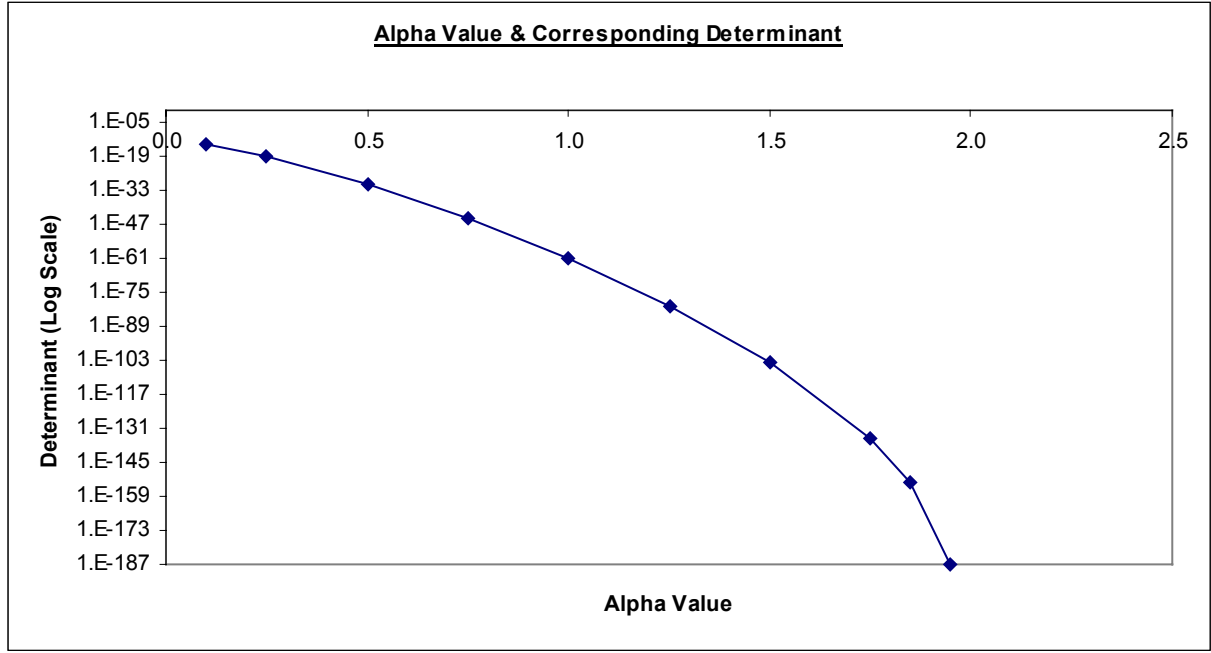


Figure 3.1: Alpha value used in semi-variogram model and determinant value of Covariance matrix

Due to the fact that the coefficient matrix is quite often very close to singular the methods of matrix inversion such as conventional Gaussian elimination and LU Decomposition do not return satisfactory results. The method of Singular Value Decomposition (SVD) is needed to determine the inverse solution of the coefficient matrix.

The method of Singular Value Decomposition involves the decomposition of the matrix that is being inverted. The resulting matrices are a column orthogonal matrix  $U$ , a diagonal matrix  $W$  which contains positive or near zero singular values and the transpose of an orthogonal matrix  $V$ . This is indicated by the equations shown below [Press et al, 1992].

$$\begin{bmatrix} G & u \\ u^T & 0 \end{bmatrix} = \begin{pmatrix} U \end{pmatrix} \cdot \begin{pmatrix} w_1 & 0 & 0 \\ 0 & w_2 & \\ 0 & \dots & w_n \end{pmatrix} \cdot \begin{pmatrix} V^T \end{pmatrix}$$

To compute the inverse one simply needs to take the inverse of the  $w_j$  values along the diagonal of the  $W$  matrix and multiply out the matrices in the order as shown below.

$$\begin{bmatrix} G & u \\ u^T & 0 \end{bmatrix}^{-1} = V \cdot \left( \text{diagonal} \left( \frac{1}{w_j} \right) \right) \cdot U^T$$

However in the instance of (near) singularity the values on the diagonal of the  $W$  matrix (the  $w_j$  values) need to be considered carefully. In some instances, as in this one, the  $w_j$  values along the diagonal can be very small but not actually zero leading to a highly ill-conditioned matrix.

As an example a coefficient matrix of size 150 by 150 was computed from a segment removed from a rainfield. As indicated in Figure 3.2 which shows the  $w_j$  values rapidly decrease in magnitude along the diagonal of the  $W$  matrix until they are very close to zero. As the value of alpha used in the semi-variogram model is increased from a value of one towards a value of two the  $w_j$  values decrease in magnitude far more rapidly and become significantly closer to zero. (Note the logarithmic scale of the vertical axis)

If the  $w_j$  values approach a value that is approximately very close to zero it is best to eliminate that particular value and replace its inverse with a zero value. This results in the elimination of values that are probably highly inaccurate due to round off errors and upon inverting will lead to highly inaccurate final solutions. This is evident when the alpha value is equal to two, where the majority of values are very nearly zero and return highly unsatisfactory answers.

The sum of the square of the  $w_j$  values along the diagonal of the matrix is also equal to the explained variance of the data points, as shown by the equation below.

$$\sum_{j=1}^n w_j^2 = \sigma_x^2$$

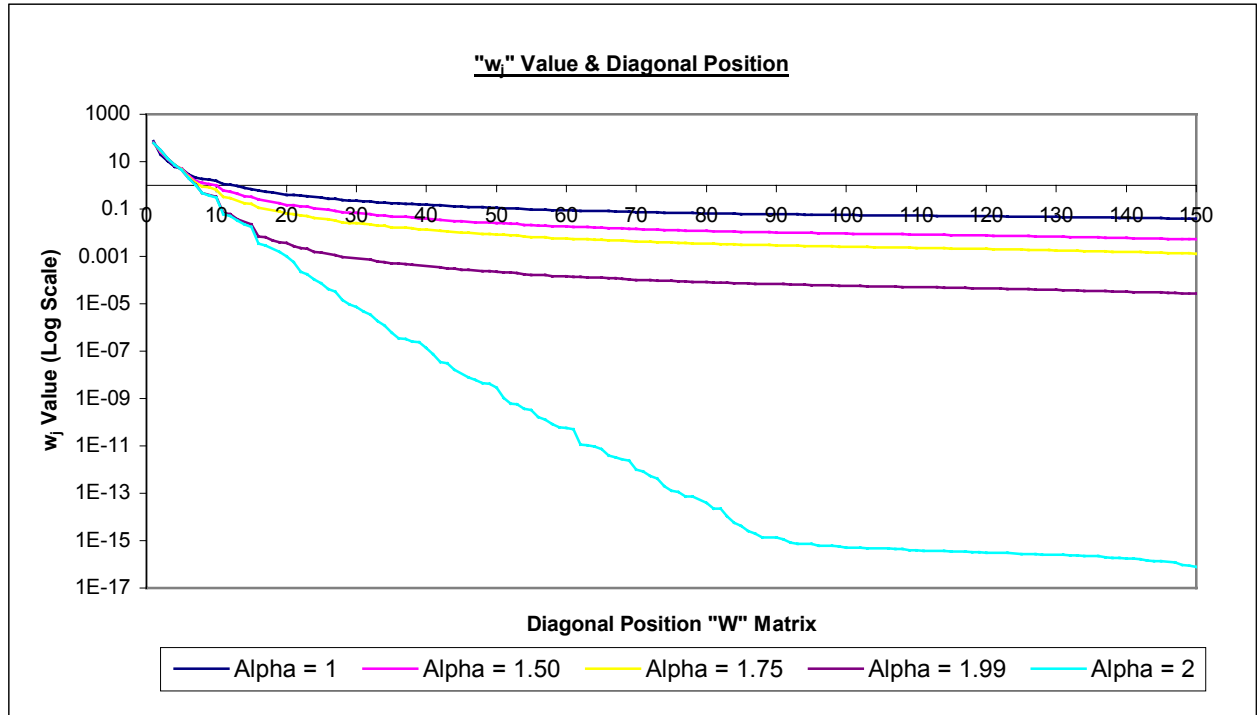


Figure 3.2:  $w_j$  value and diagonal position in W matrix for different alpha values

In most instances the  $w_j$  values are initially very high and tail off rapidly to values that are approximately close to zero. The explained variance is then given by the first few initial terms in the W matrix. As indicated in Figure 3.2 the sum of the squares initial twenty to thirty values are approximately equal to the explained variance with the remainder of the terms contributing very little to this.

To avoid the matrix being ill-conditioned and also due to the fact that the sum of squares of the initial terms are approximately equal to the explained variance it is appropriate to trim the  $w_j$  values at a certain point where they are close to zero and replace them with zero values. This not only stops the matrix being ill-conditioned but also has the added advantage that the computation time of the matrix multiplication to get the estimates can be dramatically improved with little loss in accuracy of final results.

Since the W matrix has now been trimmed of the  $w_j$  values that are approximately close to zero and replaced by zero only a few of the rows and columns of the  $U^T$  and V matrices are needed to return the solution of the inverse. The inversion process is speeded up dramatically when this is taken into account.



As an example, a segment of a rainfield was chosen and various portions of data removed in order to simulate a scenario where there is an existence of ground clutter. The missing data when infilled using Ordinary Kriging and the coefficient matrix (of size 240 by 240) inverted by Singular Value Decomposition. A Correlation Length of  $L = 11.5$  and an  $\alpha = 1$  was used in the semi-variogram model. The sum of the absolute value of the errors was then computed by comparison of the original known data with the estimated Ordinary Kriged data. This was done several times but in each case the  $W$  matrix values were trimmed along with the corresponding row and columns of the  $U^T$  and  $V$  matrix. The sum of the absolute value of the errors and the increase in computation time with the corresponding percentage of  $w_j$  values removed is indicated in Figure 3.3 and Figure 3.4.

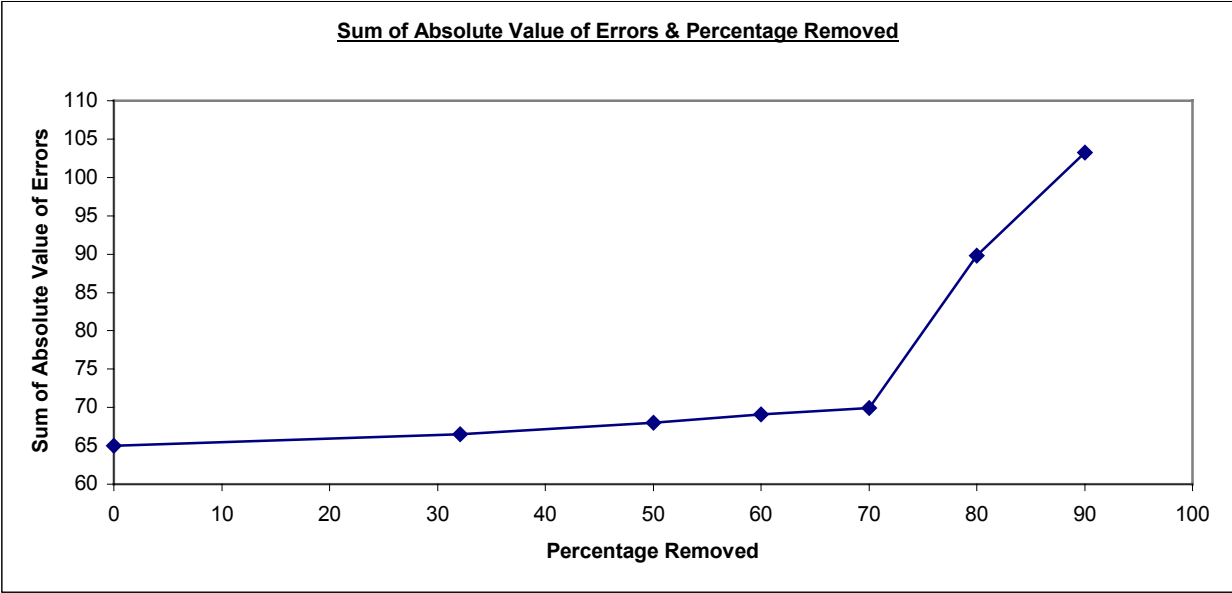


Figure 3.3: Sum of the absolute value of the errors and Percentage of  $w_j$  removed

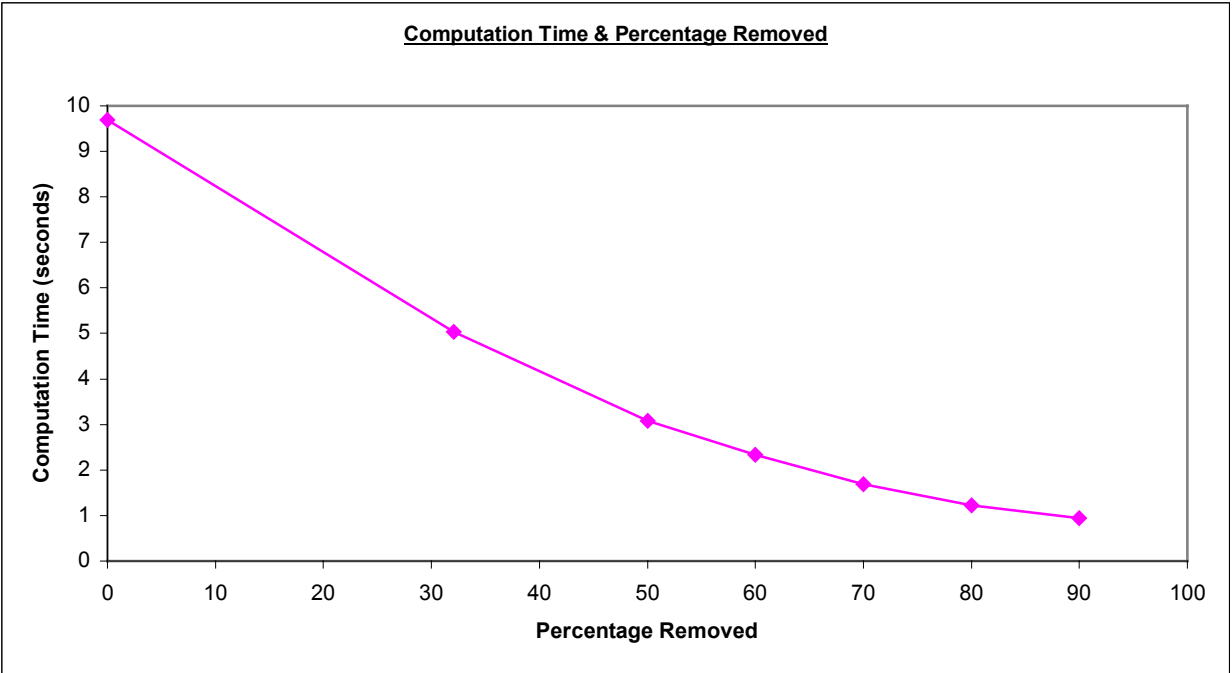


Figure 3.4: Computation time (seconds) and Percentage  $w_j$  removed

As indicated in this example just over 70 % of the  $w_j$  values could be removed until there was a noticeable decrease in accuracy of results. There is also a corresponding rapid decrease in computation time once the  $w_j$  values are trimmed as indicated by Figure 3.4. Once 70 % has been trimmed a 5 – fold increase in speed is obtained for the entire Ordinary Kriging process.

In order for Kriging to be computational fast and efficient, and for this technique to be feasible as a real time application, computationally fast and efficient methods need to be investigated for solving the sets of linear equations involved in this process.

#### **4. SUMMARY**

Radar rainfall data are of use to many organisations however there are inaccuracies associated with these data such as ground clutter and anomalous propagation to name just two. The contaminated or missing data can be estimated by the Kriging technique, which has its relative advantages and disadvantages as discussed. In order for the Kriging to be used in a real time application the process needs to be computationally fast and efficient. The distribution of weightings around the missing data allows the five to ten nearest data points to be used instead of the entire data set, which lessens the computational load significantly.

In addition, the coefficient matrix computed in the Kriging process can be highly ill-conditioned depending on the semi-variogram parameters used and the size of the matrix. Singular Value Decomposition can be used to solve this problem in conjunction with the elimination of the  $w_j$  values that are approximately close to zero. By the elimination of these values the computational speed and efficiency can also be greatly improved.

## 5. REFERENCES

Chiles, J.P. Delfiner, P. (1999). *Geostatistics – Modelling Spatial Uncertainty*. Wiley-Interscience Publication. pp 202 - 206

Cressie, N (1993). *Statistics for Spatial Data*. Wiley-Interscience Publication. pp 122 -123

Pegram, G.G.S (2001). *2-D Surface Fitting Using FFT and Numerical Convolution with Kriging and Multiquadrics*. University of Natal, Internal Report. pp 5

Press, W.H. Teukolsky, S.A. Vetterling, W.T. Flannery, B.P. (1992). *Numerical Recipes in C – The Art of Scientific Computing*. 2<sup>nd</sup> Edition. Cambridge University Press. Cambridge. pp 59, 63

Seed, A.W. Pegram, G.G.S. (2001). Using Kriging to Infill Gaps in Radar Data due to Ground Clutter in Real-Time. *Fifth International Symposium on Hydrological Applications of Weather Radar – Radar Hydrology – Proceedings*. Kyoto. pp. 73 – 78