

## Improved radar rainfall estimation at ground level

S. M. Wesson and G. G. S. Pegram

Civil Engineering, University of KwaZulu-Natal, Howard College Campus, Durban, South Africa

Received: 12 September 2005 – Revised: 23 January 2006 – Accepted: 19 February 2006 – Published: 10 May 2006

**Abstract.** A technique has been developed to provide an estimate of the rainfall reaching the earth's surface by extrapolating radar data contained aloft to ground level, simultaneously estimating unknown data in the radar volume scan. The technique has been developed so as to be computationally fast, to work in real time and comprises the following steps. A rainfall classification algorithm is applied to separate the rainfall into two separate types: convective and stratiform rainfall. Climatological semivariograms based on the rainfall type are then defined and justified by testing, which result in a fast and effective means of determining the semivariogram parameters anywhere in the radar volume scan. Then, extrapolations to ground level are computed by utilising 3-D Universal and Ordinary Cascade Kriging; computational efficiency and stability in Kriging are ensured by using a nearest neighbours approach and a Singular Value Decomposition (SVD) matrix rank reduction technique. To validate the proposed technique, a statistical comparison between the temporally accumulated radar estimates and the Block Kriged raingauge estimates is carried out over matching areas, for selected rainfall events, to determine the quality of the rainfall estimates at ground level.

### 1 Introduction

Raingauges are the traditional tool used for the recording of rainfall and are often regarded as the “true”, or reference, rainfall estimates at ground level. Raingauges have the advantage of being relatively cheap and easy to maintain and also of providing a direct estimate of the accumulated rainfall at a particular point. However there are various disadvantages associated with raingauges. They tend to underestimate during heavy rainfall periods (Wilson and Brandes, 1979)

Correspondence to: G. G. S. Pegram  
([pegram@ukzn.ac.za](mailto:pegram@ukzn.ac.za))

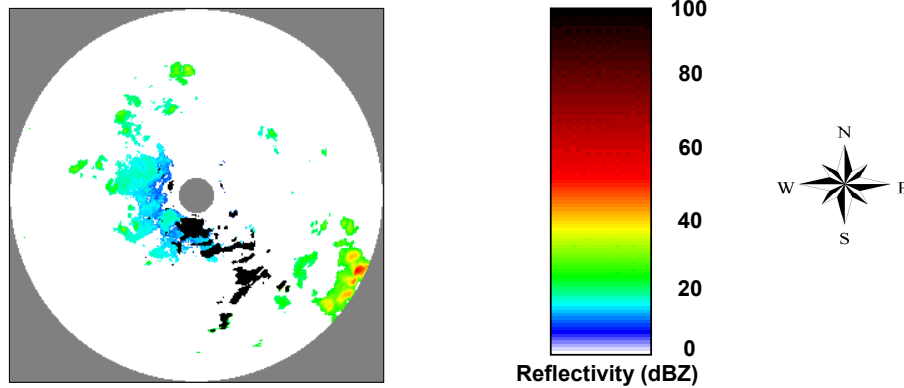
and by providing a point estimate they can fail to capture the spatial variability of rainfall. Random errors are also present which can be demonstrated by the difference in measurement obtained between closely situated raingauges (Ciach, 2002).

Weather radars overcome some of the disadvantages associated with raingauges. Currently in South Africa the weather radar network of eleven C-band radars and one S-band radar provides images of the instantaneous reflectivity values at approximately five-minute intervals and at a one-kilometre horizontal resolution supplied in Cartesian co-ordinates on Constant Altitude Plan Position Indicators (CAPPIs) at 1 km intervals above ground level. This type of data provides a detailed spatial representation of the rainfield in real time and over a large area. Figure 1 illustrates a typical instantaneous radar reflectivity image taken from the Bethlehem (South Africa) S-band weather radar where the white portions indicate no rainfall, grey portions indicate areas out of the weather radar's range (and also where there are no data available), and the black portions mark ground clutter locations.

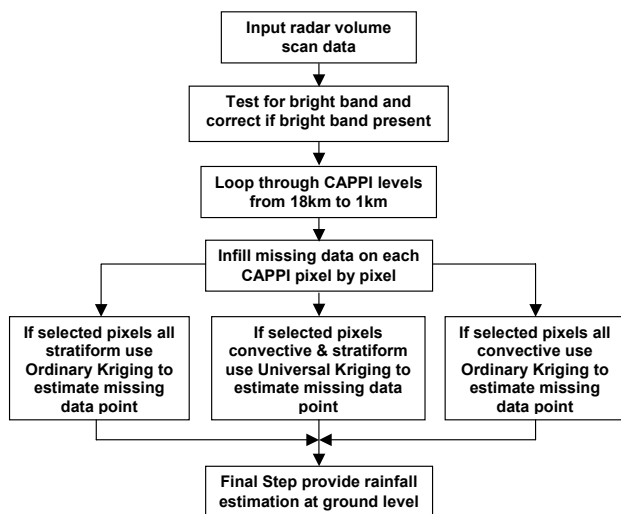
Rainfields estimated from weather radars experience various data quality problems such as ground clutter, anomalous propagation and beam blocking, to name a few (Terblanche et al., 2001). Another disadvantage of weather radars is that they provide an indirect measurement of precipitation intensity, so the returned power of measured reflectivity values has been converted to rainrate by an appropriate transformation, such as the Marshall Palmer formula (Marshall and Palmer, 1948) given by Eq. (1), where  $Z$  is the reflectivity (dBZ) and  $R$  is the rainrate (mm/h).

$$Z = 200 \cdot R^{1.6} \quad (1)$$

The data in the CAPPIs are only available from one-kilometre above ground level and there are regions within the volume scan where the rainrate is unknown. In applications, such as disaster management, hydrology and agriculture, the rainfall estimates at ground level are of more interest



**Fig. 1.** Typical radar reflectivity image taken from the Bethlehem (South Africa) weather radar on the 25 February 2003. The grey portions of the image indicate where no data are available and the black portions indicate ground clutter. The image is 300 km square.



**Fig. 2.** Simplified flow chart of the process to progressively infill all missing data in a radar volume scan and provide an estimate of the rainfall at ground level.

and importance than the measurements aloft, which are unlikely to be an accurate indication of the rainfall at ground level (Jordan et al., 2000) since precipitation tends to be affected by a variety of atmospheric phenomena before reaching the earth's surface.

Attempts have been made to provide an improved rainfall estimate at ground level by taking into account atmospheric factors such as wind drift (Lack and Fox, 2004). A different approach, taken by Franco et al. (2002), was to compute a Mean Apparent Vertical Profile of Reflectivity and by extending the profile shape to obtain an estimate of the rainfall at ground level. Currently in South Africa the real-time radar rainfall estimates at ground level are simply taken as the average vertical reflectivity in a column (Visser, 2003). Seed and Pegram (2001) proposed using 3-D Kriging to estimate

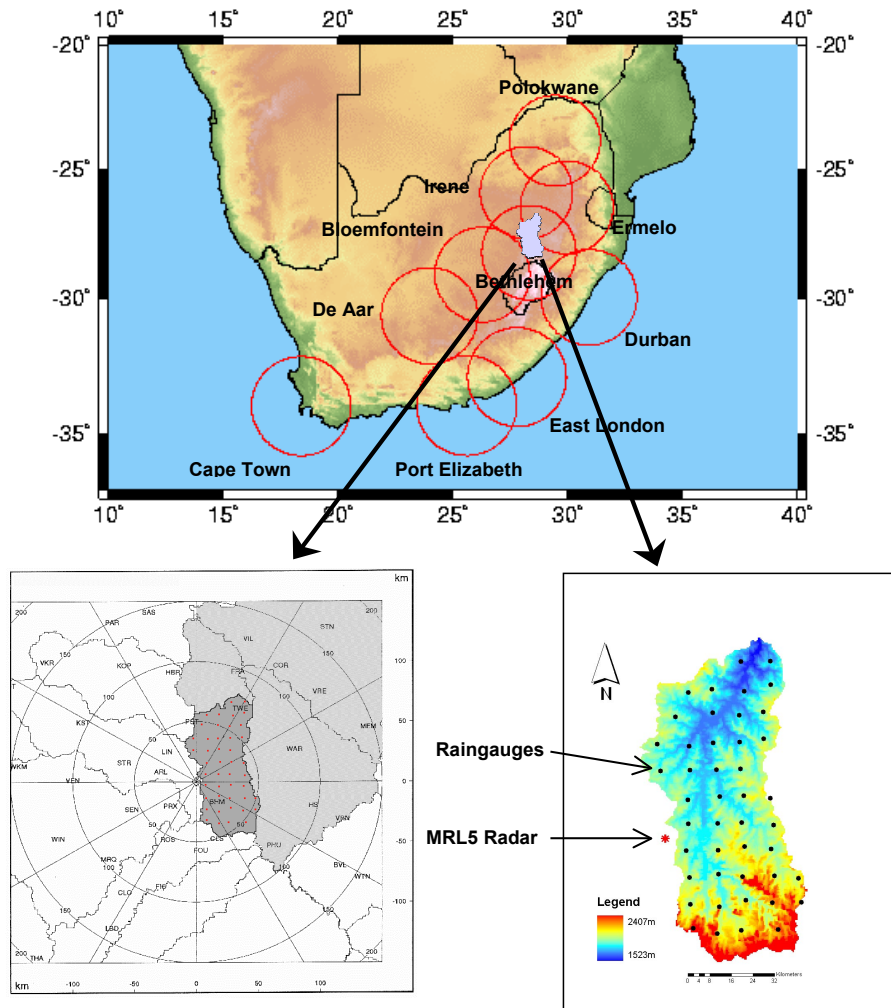
the rainfall at ground level and that paper is the source of some of the ideas for this paper.

The algorithm proposed in this paper to infill all missing data in a radar volume scan and provide an estimate of the rainfall at ground level comprises the following steps, which are also illustrated in a simplified flow chart given by Fig. 2.

- The reflectivity radar volume scan data is thresholded and the remainder of the reflectivity values classified as convective or stratiform rainfall in a pixel-by-pixel manner.
- Precalibrated climatological semivariograms based on the classified rainfall type are then defined.
- The next part of the algorithm employs a bright band correction algorithm which, where necessary, adjusts the reflectivity values of the stratiform rainfall on the 2 km CAPPI affected by bright band in a pixel-by-pixel manner.
- The reflectivity values are then extrapolated to ground level via a nearest neighbourhood Cascade Kriging approach. All the missing data on each CAPPI are infilled, working progressively from the uppermost CAPPI to ground level, where infilled data on upper CAPPIs are used in the estimation of the data on lower CAPPIs.

To determine the quality of the rainfall estimates at ground level a statistical comparison between the temporally accumulated radar estimates and the Block Kriged raingauge estimates is carried out over matching areas, for selected rainfall events.

The data from 1995 and 1996 from the Liebenbergsvlei catchment near Bethlehem were selected as the test-set for this study. This was because there was a dense network of 45 tipping bucket raingauges placed in a grid, evenly spaced at approximately 10 km throughout the catchment at that time (since decommissioned). The catchment is approximately



**Fig. 3.** Map of South Africa illustrating the weather radar network coverage and the geographic location of the Liebenbergsvlei catchment. An enlarged DEM image of the catchment shows the approximate positions of the raingauges and the MRL5 weather radar.

3600 km<sup>2</sup> in area, which gives a raingauge density of 1 raingauge per 80 km<sup>2</sup>. The topography is one where the elevation is approximately 2400 m above sea level in the Southern reaches of the catchment and slopes downwards to the Northern region of the catchment to an elevation of approximately 1500 m. The Bethlehem S-band radar is also situated in close proximity to the catchment. Both of these factors are important in ensuring accurate estimates of rainfall accumulation values. The geographic position of the area and layout of the tipping bucket raingauges are illustrated in Fig. 3 as well as a Digital Elevation Map depicting the topography of the catchment.

## 2 Rainfall classification

The initial step, of the operational algorithm for infilling the missing or inaccurate data on a CAPPI, requires a simple

classification and thresholding of the instantaneous radar volume scan data, pixel by pixel. A variety of rainfall classification algorithms have been suggested that involve subdividing the rainfall into three zones – convective, intermediate and stratiform rain. An example of a current rainfall classification algorithm employed by The Group of Applied Research on Hydrometeorology (GRAHI) (Sempere-Torres et al., 2000) also subdivides the rainfield into the same three separate groupings; the convective rain is identified by applying a threshold value of 40 dBZ; the stratiform rain is identified when a bright band is present; the rest of the rainfield is classified as transitional. Another rainfall classification algorithm developed by Mittermaier (1999) was also designed along similar principles.

The above-mentioned algorithms are not suitable for the purpose of rainfall estimation at ground level over large areas. This is due to the fact that the data at the 2 km CAPPI (which, because the base scan is at 1.5° above horizontal has

**Table 1.** Rainfall classification and threshold criteria to separate radar volume scan data into stratiform and convective rainfall domains.

Criteria	Classification
Pixel $\leq$ 18 dBZ	No Rainfall: Pixel set to 0 dBZ
18 dBZ < Pixel < 35 dBZ	Stratiform Rain
Pixel $\geq$ 35 dBZ	Convective Rain

a range limited to 73 km) plays a pivotal role in classifying the rainfall as stratiform or intermediate. This means that only a small proportion of the total volume scan data can be classified. Another drawback of the existing algorithms in real time applications is that they are computationally expensive, especially for large data sets.

The algorithm developed here, with assistance of meteorologists from the South African Weather Services (SAWS) Innovation and Research Division, classifies the rainfall into only two groups: convective and stratiform and works as follows. The classification is rather like Steiner's (1995) algorithm and is done pixel by pixel throughout the volume scan data. The initial step of the algorithm works on a simple threshold where all reflectivity values less than or equal to 18 dBZ are set to zero due to the fact that 18 dBZ approximates to 0.50 mm/h, a rate that can be considered negligible. The remainder of the reflectivity values above zero are then classified by the simple criterion: if the reflectivity value is 35 dBZ or above, the rainfall is classified as convective; the reflectivity values in the range 18 dBZ to 35 dBZ are classified as stratiform rainfall.

A summary of the classification criteria is given in Table 1 and an example of a classified instantaneous image 4 km above ground level is illustrated in the top panels of Fig. 4. An example of a vertical cross section through a volume scan data with corresponding is also illustrated in the bottom image of Fig. 4.

The classification algorithms proposed by Steiner (1995) and Mittermaier (1999) work on the basis that the classification of convective rainfall on the 4 km CAPPI is applied to all levels below and that the 2 km CAPPI is used to classify stratiform rain. However the rainrates at a large number of pixels outside the range of the 4 km and 2 km CAPPI need to be estimated and large portions of the image on the 4 km and 2 km CAPPI may also be marked as ground clutter and therefore be contaminated. As indicated in the bottom image of Fig. 4, where the convective rainfall approaches ground level (below 4 km in altitude) the convective rain zones may increase substantially in the horizontal direction, which would then result in a scenario where rainfall of high intensity (e.g. 54 dBZ  $\approx$  84 mm/h) is miss-classified as stratiform rainfall. For the above reasons a pixel by pixel classification approach was chosen throughout the CAPPI volume scan due to its

greater suitability for the purpose of rainfall estimation at ground level.

The main advantage of this algorithm is its simplicity, in that for a real time application it has little impact on the overall computation time; this is especially true for large data sets such as those returned from weather radar volume scans. Secondly the algorithm is not restricted to the 2 km CAPPI which is frequently the level of the melting layer in a South African context. Thirdly, a far greater spatial range can be classified than that limited to 73 km at the 2 km level, admittedly with reduced precision and confidence as the range increases.

## 2.1 Characteristics of classified rainfall

The rainfall classification algorithm was tested on numerous images ranging over five different years (1995, 1996, 2000, 2001 and 2002) from two weather radars in South Africa: Bethlehem (mixed convective and stratiform rain, 1800 m altitude, sub-tropical savannah) and Durban (mostly orographic warm rain, coastal, bush).

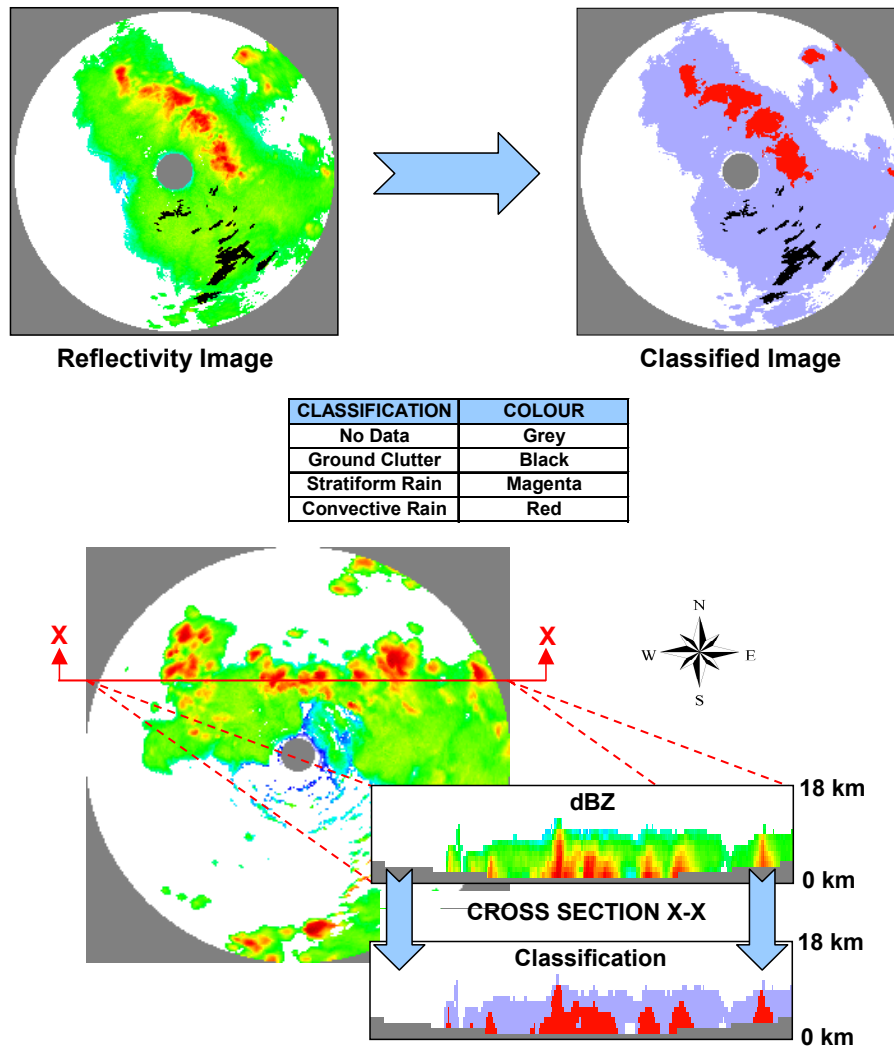
The vertical characteristics of the classified stratiform and convective rain were examined by computing the mean vertical profile of reflectivity from selected images from the Durban and Bethlehem weather radars over the same five years as mentioned above. As shown in Fig. 5 the stratiform rain generally has a low average vertical height, limited to 8 km above ground level. The convective rainfall has considerable vertical extent ranging up to 14 km above ground level and a bigger range in mean reflectivity values. Both the stratiform and convective profiles show a constant increase in rainfall intensity as ground level is approached, with the exception that in the stratiform profile, evidence of a bright band at the 2 km level indicates that the stratiform rain is likely to have been classified correctly. The variability of the reflectivity in both are characterised by a standard deviation of 3.5 dB or less at any level.

## 3 Semivariogram estimation

One of the advantages of Kriging compared to other interpolation/extrapolation techniques, is that the basis function is determined by the actual data set and takes into account the spatial structure of the data. The basis function in Kriging is determined by the computation and fitting of an appropriate model to the computed semivariogram. The variogram is defined as the expectation of the square of the differences of the field variables separated by a specified distance, as given by Eq. (2):

$$2 \cdot \gamma(x, h) = E \left\{ [Z(x) - Z(x + h)]^2 \right\} \quad (2)$$

where the semivariogram is defined as  $\gamma(x, h)$  (Journel and Huigbregts, 1978: 11). The semivariogram provides a way of measuring the spatial dependence that exists amongst the



**Fig. 4.** The top image is a radar reflectivity image from the Bethlehem weather radar (30 December 2001) and the corresponding classified image. The bottom image is a cross sectional profile through an instantaneous reflectivity image. Indicated in the associated plan is the CAPPI 4 km above ground level and the position of the vertical cross section in the radar volume scan CAPPI (24 January 2002). All images are 300 km across. Note the effectiveness of the classification algorithm.

variables in a stationary random field. The Kriging computation can be carried out using either a covariance or a semivariogram function. However in this application the semivariogram is used in preference due to its more robust properties as outlined by Cressie (1993: 70–73).

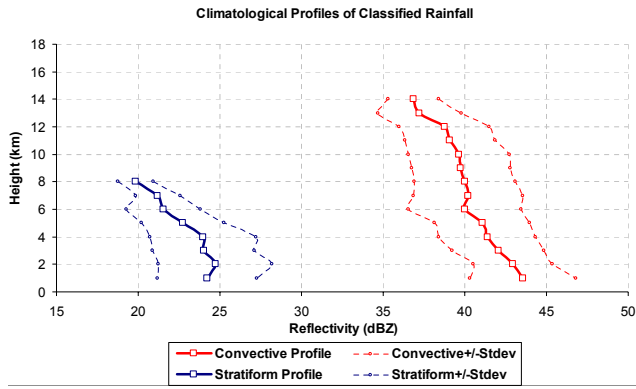
### 3.1 Empirical computation of the semivariogram

The semivariogram is most commonly computed by the Classical Variogram estimator as was proposed by Matheron (1962) and is given by Eq. (3):

$$2 \cdot \hat{\gamma}(h) = \frac{1}{|N(h)|} \sum_{N(h)} [Z(s_i) - Z(s_j)]^2 \quad (3)$$

where  $\hat{\gamma}(h)$  is the sample semivariogram at lag  $h$  (the specified lag distance),  $Z(s_i)$  and  $Z(s_j)$  are the values of the variables at the specified locations,  $s_i$  and  $s_j$  which are  $h$  apart and  $N(h)$  is the number of pairs separated by lag  $h$ .

Unfortunately the classical estimator of Eq. (3) is badly affected by non-typical observations which can be attributed to the  $(\cdot)^2$  term in the summand (Cressie, 1993: 40). The effect of non-typical data points can have a dramatic effect on the semivariogram since they are used numerous times in the calculation at different lag intervals. This can result in peaks or shifting of the entire semivariogram upwards (Sabyasachi et al., 1997). It was also noted that the distribution of points at each lag distance ( $h$ ) computed from radar rain fields was highly skewed and did not approximate a normal distribution.



**Fig. 5.** Average vertical profile of reflectivity for classified stratiform and convective rainfall. Profiles are computed from 20 different images ranging over a 5 year period from the Bethlehem and Durban weather radars.

It was thus decided that the Robust Variogram, proposed by Cressie and Hawkins (1980) was the more appropriate model to use and is given by Eq. (4):

$$2 \cdot \bar{\gamma}(h) = \left\{ \frac{1}{|N(h)|} \sum_{N(h)} |Z(s_i) - Z(s_j)|^{1/2} \right\}^4 / \left( 0.457 + \frac{0.494}{|N(h)|} \right) \quad (4)$$

By computing the sum of the absolute difference of the pairs of data points,  $Z(s_i)$  and  $Z(s_j)$  in the square root domain and then raising the result to the power of four dramatically reduces the effect of uncharacteristic observations. It was found that the skewness of the points at each lag ( $h$ ) was greatly reduced and also that the mean and the median value of the sets of points for each lag interval more closely coincided, both indicating that the distribution of points at each lag ( $h$ ) more closely approximates a normalized distribution.

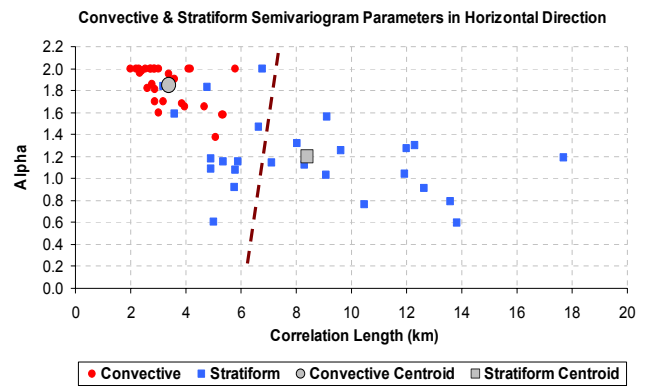
The two-parameter isotropic exponential model was chosen as the model to fit to the empirical semivariogram values, as defined by Eq. (5):

$$g(h) = 1 - \exp[-(h/L)^\alpha] \quad (5)$$

where  $h$  is the Euclidian distance between data points,  $L$  the correlation length and  $\alpha$  the shape parameter which lies in the range  $0 < \alpha \leq 2$ . As recommended by Journel and Huijbrechts (1978: 194), the model was fitted using up to only half the maximum possible lag and only to the bins at each lag distance that contained thirty or more points.

### 3.2 Parameter fitting to rainfall types

Computing the empirical semivariogram values and then applying a model fitting routine to solve for the  $\alpha$  and  $L$  parameters in Eq. (5) is a computationally burdensome and time-consuming task, especially if this is done each time an un-



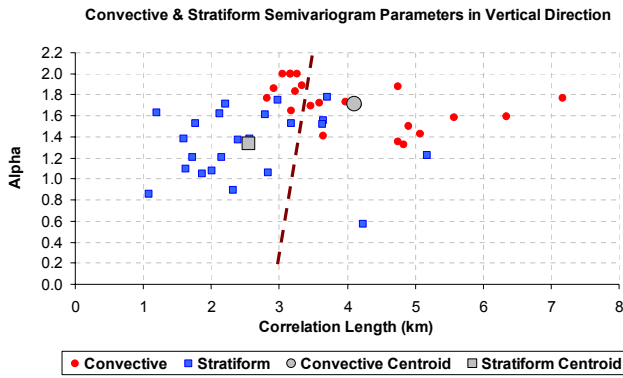
**Fig. 6.** Scatter plot of alpha ( $\alpha$ ) and correlation length ( $L$ ) parameters for stratiform and convective rainfall in a horizontal direction. The dashed line indicates the separation boundary of the two clusters as identified by a Fuzzy C-Means cluster algorithm.

known reflectivity value is estimated. This problem is exacerbated in the context of real time calculation, when the images are collected every 5 min. An alternative is therefore needed that is computationally efficient yet does not compromise the accuracy of the estimates. Reducing the computational burden of computing the semivariogram by the use of a climatological semivariogram was employed for raingauge interpolation by Lebel et al. (1987) in order to improve computational efficiency. This idea was explored in the context of radar rainfields.

A set of instantaneous rainfall images was selected from the Bethlehem and Durban weather radars ranging over a period of several years (1995, 2000, 2001 and 2002). The images consisted of a variety of rainfall types ranging from images containing solely stratiform rain to images containing a combination of both stratiform and convective rainfall.

The rainfall classification algorithm of Table 1 was applied to the CAPPI 4 km above ground level. From the images various rectangular portions were selected that contained either predominantly stratiform or convective rainfall. These regions were approximately 1200 km<sup>2</sup> (40 km by 30 km) in area or larger where the reflectivity data was standardised before computing the empirical semivariogram values. From the selected regions the Robust semivariogram was computed and the two parameter exponential model (Eq. 5) fitted. This was done separately in both the horizontal and vertical directions. The two parameters,  $\alpha$  and  $L$  estimated from each field, were then plotted against one another in a scatter plot. This is shown in Fig. 6 for the horizontal direction and Fig. 7 for the vertical direction.

The scatter plots in Figs. 6 and 7 illustrate that the semivariogram parameters tend to cluster in a particular region according to the type of rainfall, indicating that a fixed set of parameters can comfortably be used depending on the rainfall type.



**Fig. 7.** Scatter plot of alpha ( $\alpha$ ) and correlation length ( $L$ ) parameters for stratiform and convective rainfall in a vertical direction. The dashed line indicates the separation boundary of the two clusters as identified by a Fuzzy C-Means cluster algorithm.

### 3.3 Cluster analysis

To test if a natural clustering of the variables occurs, and also to ascertain the degree of clustering in the vertical and horizontal directions, a cluster analysis technique was utilised. A Fuzzy C-Means clustering algorithm (Gordon, 1981: 58–60) was run on the scatter plot of the data, where the specification set in the algorithm was that the data should be divided into two groups. In the horizontal direction the Fuzzy C-means cluster algorithm split the groups roughly along the 6.5-km correlation length and in the vertical direction along the 3.5-km correlation length. The divisions for the clusters in the vertical and horizontal direction determined by the Fuzzy C-means cluster algorithm are indicated on Figs. 6 and 7 by the dashed line.

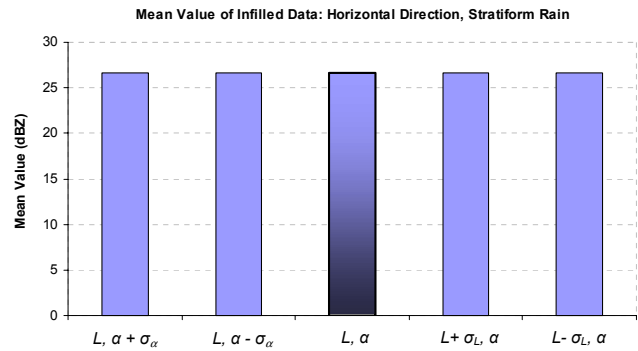
Because the semivariogram parameters naturally cluster around a centroid depending on the rainfall type, it seems to be a safe assumption to use fixed climatological semivariogram parameters depending on the type of rainfall classified; the adopted values are given in Table 2. This idea is tested in the next sub-section.

### 3.4 Sensitivity analysis of climatological semivariogram parameters

To determine how sensitive final Kriged solutions are to the use of a fixed set of semivariogram parameters, a sensitivity analysis was undertaken. In the testing procedure, various 2-D data sets of radar reflectivity values were selected that contained solely convective or stratiform rainfall. Random portions of the selected reflectivity data were then removed and Ordinary Kriging was used to estimate the missing data, but in this instance the Ordinary Kriging was carried out several times using different combinations of semivariogram parameters (a brief overview of Kriging methods will be given in a later section). Five different sets of parameters were used: firstly the centroid value of  $\alpha$  and  $L$  for convective and strati-

**Table 2.** Semivariogram parameter values in the horizontal and vertical directions for stratiform and convective rainfall.

	Horizontal		Vertical	
	$\alpha_H$	$L_H$ (km)	$\alpha_V$	$L_V$ (km)
Stratiform	1.53	8.40	1.33	2.56
Convective	1.85	3.38	1.71	4.11



**Fig. 8.** Comparison of the mean value for an infilled region with five different combinations of alpha ( $\alpha$ ) and correlation length ( $L$ ) parameters as indicated on the horizontal axis on the graph.

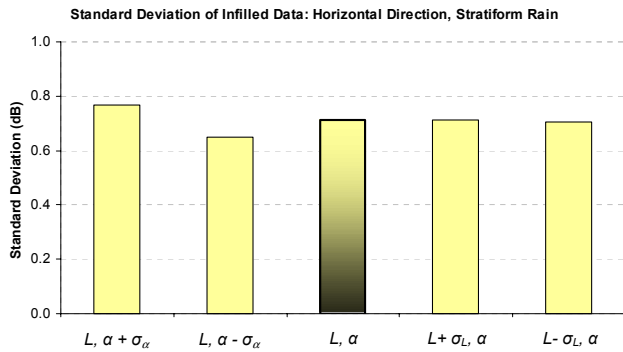
form rain and then the  $\alpha$  and  $L$  values one standard deviation away from the centroid values, as computed by Eqs. (6) and (7).

$$\alpha \pm \sigma_\alpha, \quad 0 < \alpha \leq 2 \tag{6}$$

$$L \pm \sigma_L, \quad 0 < L \tag{7}$$

The infilled data were then compared to the original observed reflectivity data by computing the Sum of Square of Errors (SSE), means and standard deviations of the observed and estimated data. As indicated by Figs. 8 and 9 there is no significant difference between the estimated values returned for the expected range of parameter values. The estimated Kriging values appear to be more sensitive to the shape parameter,  $\alpha$ , than the correlation length,  $L$ .

The results indicate that there is no significant difference between the five sets of parameter values used in each instance, giving a clear indication that the final Kriged solution is reasonably insensitive to the range of the shape parameter,  $\alpha$ , and correlation length,  $L$ , values that can possibly occur when a sample semivariogram is fitted to a specific rainfall type. Instead of solving for each set of semivariogram parameters for each neighbourhood one can simply use the centroid value computed for each rainfall type depending on the observed rainfall.



**Fig. 9.** Comparison of the standard deviation for an infilled region with five different combinations of alpha ( $\alpha$ ) and correlation length ( $L$ ) parameters as indicated on the horizontal axis on the graph.

### 3.5 Application to 3-D reflectivity data

When selecting controls for computing the estimated value of a target in 3-D space, three possible combinations of points can be selected: (1) all convective, (2) mixed, containing both stratiform and convective points and (3) all stratiform. For situations with either all convective or all stratiform, the appropriate subset of parameters in Table 2 are utilised for all estimations.

When it comes to the practical calculation of the semivariogram elements in the Kriging equations (Eq. 15) there are no numerical or technical difficulties if the controls are all of one type, either convective, or stratiform. A problem arises when the controls are of mixed type. It was found that if the individual semivariogram elements in the equation were set to their type, with a compromise value for a mixed pair, unpredictable instabilities occurred in the solution of the equation. The methodology adopted was to use a linear weighting of the  $\alpha$  and  $L$  parameters, where the weights are the proportion of type.

When Kriging with a 3-D data set, the semivariogram model developed by Seed and Pegram (2001) was modified and is given by Eq. (8):

$$\gamma(h) = \sigma^2 \cdot [1 - \exp(-h^\alpha)] \quad (8)$$

where:  $h^2 = \left(\frac{r}{L_H}\right)^2 + \left(\frac{z}{L_V}\right)^2$  and  $r^2 = x^2 + y^2$  here  $\sigma^2$  is the field variance,  $h$  is the scaled distance in spherical coordinates,  $\alpha$  is the scaling exponent,  $r$  is the distance in the horizontal plane,  $L_H$  the horizontal correlation length,  $z$  is the distance in the vertical direction,  $L_V$  the vertical correlation length and  $x$  and  $y$  are the distances in the Cartesian horizontal direction.

## 4 Bright band correction

As snow and ice crystals drop through the 0°C isotherm they begin to melt and are surrounded by a thin layer of water.

The result is that, from the radar's point of view, the melting snow and ice crystals resemble large blobs of water and the reflectivity values for the melting level are enhanced greatly, resulting in an overestimation of rain (Sanchez-Diezma et al., 2000). This level is called the bright band. At Bethlehem, South Africa the bright band generally occurs 2 km above ground level (Mittermaier, 1999) as can be seen by the stratiform climatological profile of reflectivity in Fig. 5.

A typical method to correct the bright band as described by Mittermaier (2003) is to derive a vertical profile based on radar volume scan data which is then used to correct the values and provide an estimate of the rainfall rate at ground level (e.g. Sanchez-Diezma et al., 2000).

### 4.1 Effect of bright band on ground level estimates

The presence of a bright band has an adverse effect on the estimates at ground level when applying an extrapolation algorithm to volume scan radar reflectivity data. The reflectivity values at the 2 km CAPPI, which are on average greater than the values on the 1 km CAPPI, create a trend of decreasing reflectivity values as ground level is approached. When extrapolating to ground level and pixels are selected from the 1 km and 2 km CAPPIs, the trend of decreasing reflectivity is extrapolated to ground level resulting in lower estimates of stratiform rainfall at ground level than should be expected. There is therefore a need to adjust the reflectivity values at the 2 km CAPPI to provide an improved estimate of the rainfall at ground level, as illustrated in Fig. 10.

### 4.2 Bright band correction algorithm

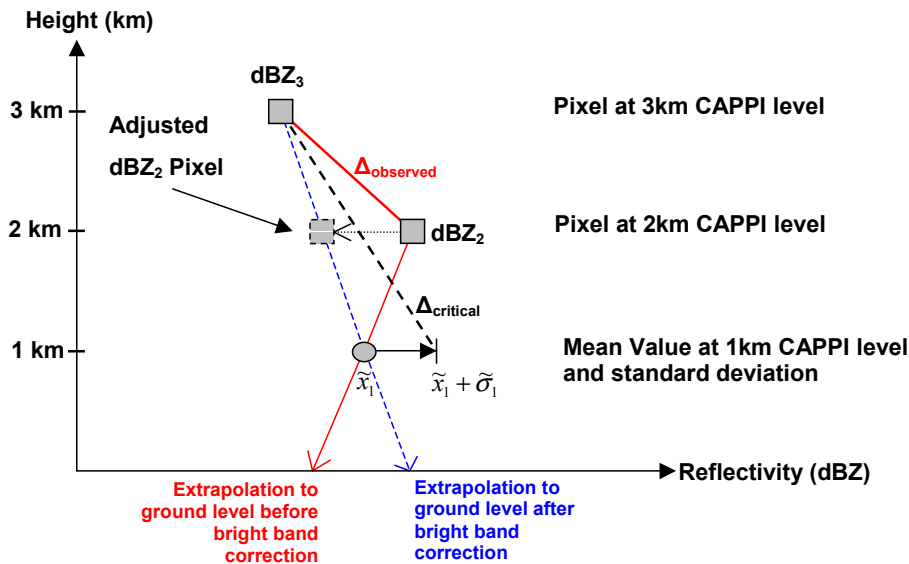
The first step of the bright band correction algorithm is to compute the wetted area ratio (WAR) of observed stratiform rainfall at the 3 km, 2 km and 1 km CAPPIs. If the WAR of stratiform rainfall is less than 10% at any of the CAPPIs, then no correction takes place due to there being insufficient rainfall to determine if a bright band is present; otherwise the 2 km CAPPI is examined to detect evidence of a bright band.

The existence of the bright band is determined by comparing the temporally weighted mean values of reflectivity ( $\bar{x}_1$ ,  $\bar{x}_2$  and  $\bar{x}_3$ ) at the 1 km, 2 km and 3 km CAPPIs; the calculation is limited to the reflectivity values classified as stratiform rainfall. If the computed temporally weighted mean values of the stratiform rainfall at each level are  $\bar{x}_1 > \bar{x}_3$  and  $\bar{x}_2 \geq \bar{x}_1$  then a bright band is considered to be present. Some memory of the mean value from previous time steps is retained by temporal weighting as indicated by Eq. (9):

$$\tilde{x}_T = \lambda \cdot \bar{x}_T + (1 - \lambda) \cdot \tilde{x}_{T-1} (\lambda = 0.50) \quad (9)$$

where  $\tilde{x}_T$  is the temporally weighted mean,  $\lambda$  represents the weighting value which is set to 0.50 (a value chosen to smooth the means, but not excessively), and  $\bar{x}_T$  is the spatial mean value of reflectivity at a given level computed from the stratiform data at time  $T$ .





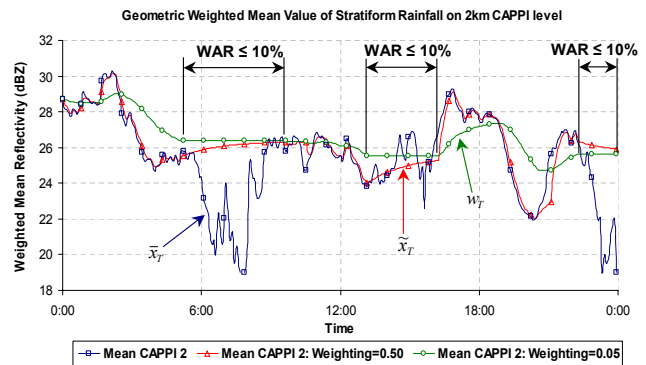
**Fig. 10.** Adjustment of a stratiform pixel value on the 2 km CAPPI that is identified as being affected by bright band. The bright band pixel is adjusted so as to sit at the midpoint between the observed reflectivity value situated on the 3 km CAPPI and the mean value of stratiform rainfall reflectivities on the 1 km CAPPI; the mean is used because the 1 km CAPPI has very limited range.

Throughout the rain event a second temporally weighted mean is also computed where the weighting value is set to 0.05; this has the result of producing a slowly adapting mean value  $w_T$ . If the WAR of the stratiform rainfall is below 10% then a break occurs and the  $\tilde{x}_T$  value is now computed based on the final  $w_T$  value before the break ( $w_{TB}$ ). An example of this is illustrated in Fig. 11 where the mean values of reflectivity ( $\tilde{x}_T$ ) are computed by geometric weighting. The reflectivity data are from the Bethlehem weather radar, 25 January 1996. As indicated in Fig. 11, in three periods, the WAR of stratiform rainfall is less than 10%; in these periods  $\tilde{x}_T$  is computed from the final  $w_T$  value before the break so that toward the end of the period, it tends to  $w_{TB}$ , which can be thought of as a locally adaptive historical mean value of reflectivity.

To correct the bright band it is necessary to compute the standard deviation ( $\sigma$ ) of the stratiform rainfall on the 1km CAPPI, as illustrated by Fig. 10. Some memory of the  $\sigma$  value from the previous time step is retained by weighting as in Eq. (9); to compute the weighted standard deviation, Eq. (10) is used:

$$\tilde{\sigma}_T = \lambda \cdot \sigma_T + (1 - \lambda) \cdot \tilde{\sigma}_{T-1} (\lambda = 0.50) \quad (10)$$

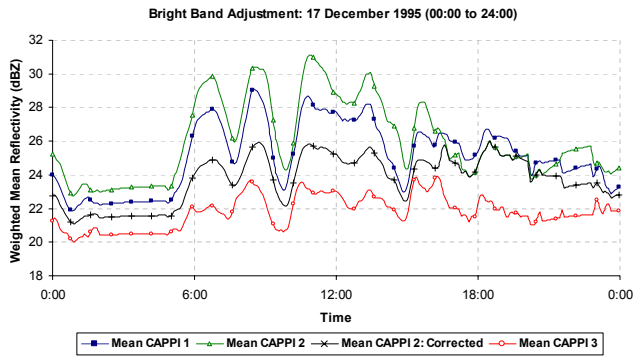
where  $\tilde{\sigma}_T$  is the weighted standard deviation,  $\lambda$  represents the weighting value which is set to 0.50 and  $\sigma_T$  is the standard deviation of reflectivity at time  $T$ . Once again throughout the rain event a second temporally weighted standard deviation ( $s_T$ ) is also computed where the weighting value is set to 0.05. If the WAR of the stratiform rainfall is below 10% then a break occurs and the  $\tilde{\sigma}_T$  value is now computed based on the final  $s_T$  value before the break ( $s_{TB}$ ).



**Fig. 11.** Weighted mean value of stratiform rain on the 2 km CAPPI as computed by geometric weighting. The data is taken from the Bethlehem weather radar (25 January 1996) over a 24-h period.

If a bright band is identified, all the values classified as stratiform rainfall are then evaluated on an individual pixel by pixel basis within a vertical column, to decide if the pixel at the 2 km CAPPI is affected by the bright band. For each stratiform pixel location on the 2 km CAPPI the gradient ( $\Delta_{critical} = \Delta \text{ dBZ} / \Delta \text{ height}$ ) is calculated between the pixel value (directly above it) on the 3 km CAPPI and the value equal to  $\tilde{x}_1 + \tilde{\sigma}_1$  on the 1 km CAPPI, as illustrated in Fig. 10; the gradient is computed from Eq. (11). The gradient ( $\Delta_{observed}$ ) is also calculated between the pixel on the 3 km and 2 km CAPPI; as indicated by Eq. (12):

$$\Delta_{critical} = \frac{\Delta \text{ dBZ}}{\Delta \text{ Height}} = \frac{\text{dBZ}_3 - (\tilde{x}_1 + \tilde{\sigma}_1)}{2} \quad (11)$$



**Fig. 12.** Bright band adjustment of the 2 km CAPPI for the 17 December 1995 (00:00 to 24:00). The green line with triangle markers indicates the 2 km CAPPI affected by bright band and the black line with plus sign markers indicates the mean value of the corrected data on the 2km CAPPI. (The data are recorded at 5-min intervals – the markers do not match the data but are for series identification in monochrome reproduction.)

$$\Delta_{\text{observed}} = \frac{\Delta \text{dBZ}}{\Delta \text{Height}} = \frac{\text{dBZ}_3 - \text{dBZ}_2}{1} \quad (12)$$

If  $|\Delta_{\text{observed}}| > |\Delta_{\text{critical}}|$  then the pixel on the 2 km CAPPI is classified as a bright band pixel and adjusted by altering the value of that pixel as shown in Eq. (13):

$$\text{New dBZ}_2 \text{ Pixel} = \frac{\text{dBZ}_3 + \tilde{x}_1}{2} \quad (13)$$

Figure 10 provides a schematic of a scenario where a pixel on the 2 km CAPPI is identified as being affected by bright band. The correction of the pixel is then also indicated as the new value is set to the midpoint of the pixel on the 3 km CAPPI and the mean value ( $\tilde{x}_1$ ) of the reflectivity data on the 1 km CAPPI.

#### 4.3 Testing the bright band correction procedure

To test the effectiveness of the proposed bright band correction procedure a period of rainfall from the Bethlehem weather radar was selected (17 December 1995) that consisted predominantly of stratiform rainfall and showed clear evidence of a bright band. Figure 12 illustrates the computed mean values  $\tilde{x}_T$  at CAPPIs 1, 2 and 3 at 5-min intervals throughout the day. The mean value of 2km CAPPI (green line with triangle markers) is consistently higher throughout the day than the mean values on the 1km and 3 km CAPPIs, indicating a definite bright band.

The bright band correction procedure described in the previous sub-section was applied to each 5-min instantaneous image throughout the 24-h period, so that the mean value sequence of the adjusted reflectivity data on the 2 km CAPPI was then recomputed. Figure 12 illustrates the adjusted mean value series for the 2 km CAPPI, which now lies at the approximate midpoint of the mean value series for the 1 km and 3 km CAPPIs.

**Table 3.** Sum of square of errors (SSE) and  $r^2$  value between radar and raingauge estimates before and after bright band correction.

	Before bright band correction	After bright band correction
SSE	16 601	13 161
$r^2$	0.34	0.43

To test the effectiveness of the bright band adjustment algorithm, a comparison was carried out between Block Kriged raingauge estimates and the radar accumulation estimates before and after the bright band correction procedure. The raingauge network utilised was from the Liebenbergsvlei catchment as described in the Introduction. The comparison was carried out on the 24-h rain event from the Bethlehem weather radar from the 17 December 1995 with a strong bright band evident, indicated in Fig. 12. The SSE between the raingauge and radar accumulations for a 24-h period were computed and the  $r^2$  values also estimated before and after the bright band correction. There was a notable improvement in the radar estimates as the radar accumulation values more closely matched the raingauge estimates after the bright band was corrected. This is demonstrated by the lower SSE and higher  $r^2$  value returned for radar estimates where the bright band correction was implemented as summarised and shown in Table 3.

The proposed bright band correction procedure provides an effective means of adjusting the 2 km CAPPI reflectivity values when affected by bright band so as to provide an improved estimate of the rainfall at ground level. The bright band correction algorithm is computationally inexpensive and therefore suitable for real time applications.

## 5 Kriging methods

Kriging was chosen as the computational method to interpolate the observations missing internal to the CAPPI stack and to extrapolate the reflectivity values contained aloft to ground level. Kriging is considered to be the optimal technique for the spatial prediction of Gaussian data (Cressie, 1993: 106). In this application two types of Kriging were used, Ordinary and Universal Kriging (also called External Drift Kriging (Hengl et al., 2003)).

### 5.1 Ordinary and Universal Kriging

In Ordinary Kriging, the mean is assumed constant and unknown throughout the field whereas in Universal Kriging, the mean is also assumed to be unknown but varying (Chiles and

Delfiner, 1999: 151). The Ordinary Kriging equations are given by Eq. (14):

$$z(s_0) = \lambda^T(s_0) \cdot z \quad (14)$$

where  $z(s_0)$  is the value to be estimated at the target  $s_0$ ,  $z$  is the vector of known reflectivity values, or controls, and the row vector  $\lambda^T(s_0)$  contains the calculated weighting values which depend on  $s_0$ . The vector of weighting values is computed by the matrix Eq. (15):

$$\begin{bmatrix} G & u \\ u^T & 0 \end{bmatrix} \cdot \begin{bmatrix} \lambda(s_0) \\ \mu(s_0) \end{bmatrix} = \begin{bmatrix} g(s_0) \\ 1 \end{bmatrix} \quad (15)$$

where  $G$  is a matrix of semivariogram values between the controls,  $u$  a unit vector of ones,  $\lambda(s_0)$  the vector of weighting values,  $\mu(s_0)$  a Lagrange multiplier and  $g(s_0)$  a vector of semivariogram values between the target at  $s_0$  and the controls' locations; note the dependence of the unknown and right hand side vector on the position of the target  $s_0$ . In Universal Kriging variables can be added to Eq. (15) in order to model the variation of the mean throughout the field. This modification is indicated in Eq. (16a), by the addition of  $Q_c$  to the coefficient matrix and  $q_t(s_0)$  to the right hand side and  $\beta$  to the vector of weighting values:

$$\begin{bmatrix} G & u & Q_c \\ u^T & 0 & 0 \\ Q_c^T & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} \lambda(s_0) \\ \mu(s_0) \\ \beta \end{bmatrix} = \begin{bmatrix} g(s_0) \\ 1 \\ q_t(s_0) \end{bmatrix} \quad (16a)$$

where the additional terms are:

$$Q_c = \begin{bmatrix} q_1(s_1) & \cdots & q_p(s_1) \\ \vdots & & \vdots \\ q_1(s_n) & \cdots & q_p(s_n) \end{bmatrix} \quad (16b)$$

The selected variable  $q_j(s_i)$  is associated with the point  $s_i$  and could be altitude, mean annual precipitation or some other classification variable. They range from  $q_1(s_i)$  to  $q_p(s_i)$  and are associated with the selected control. The vector  $q_t(s_0)$  contains the corresponding variables associated with the target. Typical models for the variation of the mean are either linear or polynomial; alternatively variables associated with physical characteristics of the controls may be included, that influence the mean value of the field.

## 5.2 Computational stability and efficiency

In order to supply an estimate of the rainfall at ground level on a real time basis, computational efficiency is of the utmost importance. In South Africa, as volume scan data are to be processed at approximately 5 min intervals, an estimated 120 000 targets need to be computed in this time. One of the disadvantages of the Kriging technique is that it relies on the solution of a linear system of equations whose size is proportional to the number of selected controls. For large systems of equations this can be time consuming, computational burdensome and potentially unstable. Previous authors have also

made reference to the computational burden associated with Kriging (e.g. Creutin and Obled, 1982).

There are several approaches that can be adopted for Kriging target data, some of which are:

- compute all target values together.
- separate the targets into contiguous sets and identify their boundary controls.
- use neighbourhood Kriging to estimate each target value individually.

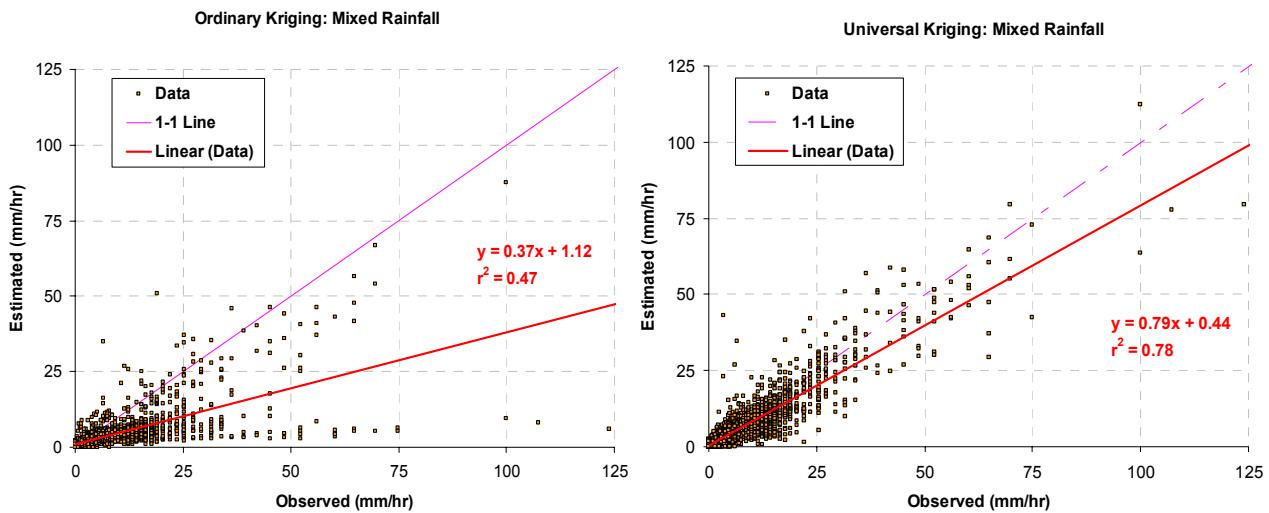
It was ascertained that the last of the three options is the most accurate, most stable and fastest of the three by an order of magnitude (Wesson and Pegram, 2004) and is the method adopted here. One of the properties of Kriging that can be exploited to overcome the problem of dimension is the “screening effect” (Chiles and Delfiner, 1999: 202–206). The screening effect refers to the observation that the significant weighting values are concentrated around the target, with the weighting values rapidly decreasing with distance from the target. For each target only a cluster of the nearest controls needs to be selected. In this application the computed optimum number of controls based on an exhaustive comparison, is 25. This has the effect of drastically reducing the computation time with little or no loss in accuracy of the final results.

Another unexpected problem associated with the Kriging technique, which is not so well known, is the ill-conditioning of the coefficient matrix. It was demonstrated by Wesson and Pegram (2004) that the coefficient matrix can be highly ill-conditioned dependent on a number of factors, but most notably the chosen parameterization of the semivariogram function. This is especially evident as the  $\alpha$  value tends to become Gaussian ( $\alpha \rightarrow 2$ ). The method of Singular Value Decomposition (SVD) is therefore used in conjunction with a trimming of the singular values to provide a computational stable method to find the solution to the Kriging equations. Due to the fact the singular values are trimmed, matrix rank reduction techniques can then be exploited to dramatically decrease the time to find the inverse solution of the coefficient matrix with little or no loss in accuracy of the final Kriged solution (Wesson and Pegram, 2004).

## 5.3 Development of Ordinary and Universal Kriging techniques

In the application of the Kriging technique the control data selected divides into three categories which are determined by the 25 controls in the selected neighbourhood and are defined as follows:

- Stratiform – controls consist entirely of stratiform rain.
- Mixed/Intermediate – controls consist of a combination of stratiform and convective rain.



**Fig. 13.** Scatter plot of observed and estimated rainrates at individual pixel points for mixed rainfall for a instantaneous image from the Bethlehem weather radar, 24 January 2002, the number of pixels estimated number 2628. Estimates are computed with Ordinary (left) and Universal Kriging (right) with the latter indicating a substantial improvement in estimates.

- Convective – controls consist entirely of convective rain.

Although the rain is classified as purely stratiform or convective rainfall an intermediate rainfall zone is implied for computation purposes. Ordinary Kriging was used in the “pure” convective and stratiform zones while Universal Kriging was used in the mixed zone. The problem to be solved was how to get the best combination in the mixed cases.

In order to determine the most effective form of Kriging to use, and the most appropriate variables to use in Universal Kriging, a validation technique was devised. The CAPPI 2 km above ground level was selected as the target level for a variety of weather types. The data from aloft was then used to estimate all the reflectivity data on the 2 km CAPPI. The estimated data and observed data were then converted from reflectivity to rainrate values and the rain separated into the three categories listed above. These were compared in terms of SSE, means and standard deviations. Universal Kriging was investigated by modelling the vertical variation of the mean by use of either a second order polynomial or a linear relationship to take advantage of the trend in the vertical reflectivity profile. This was investigated for stratiform, convective and mixed rainfall.

### 5.3.1 Stratiform zone

In the stratiform zone Ordinary Kriging is utilised. Universal Kriging did not result in any significant improvement in terms of the SSE between the estimated and observed rainfall. A computational advantage of Kriging in the stratiform zone is that SVD does not need to be used in order to compute the inverse of the coefficient matrix. This is due to the

fact that the shape parameter,  $\alpha$ , for stratiform rain is close to unity (implying an exponential semivariogram model) which results in a stable coefficient matrix. By computing the condition number, which for a symmetric matrix is defined as the ratio of the largest to the smallest singular value (Wilkinson, 1988: 191), it was shown that the coefficient matrices for stratiform rain were stable enough to use the quicker LU decomposition algorithm to find the solution to the Kriging equations.

### 5.3.2 Convective zone

The controls in this zone consist entirely of convective rainfall, so SVD is used to compute the inverse of the coefficient matrix because the shape parameter,  $\alpha$ , is close to Gaussian. Universal Kriging was not found to provide a significant improvement in the estimates over Ordinary Kriging. When applying Universal Kriging the linear relationship tended to overestimate the convective rainfall and the second order polynomial provided an underestimate at the target. Based on the above, Ordinary Kriging is used when the controls selected consist entirely of convective pixels.

### 5.3.3 Mixed/intermediate zone

In the mixed zone, SVD is used to compute the inverse of the coefficient matrix because, for convective rain, the shape parameter  $\alpha$  is close to Gaussian as seen in Table 2. This results in a numerical unstable coefficient matrix (Wesson and Pegram, 2004). In addition Universal Kriging in this instance provides an improved estimate over Ordinary Kriging. Testing the method on 10 different instantaneous images showed an improvement in estimates (in terms of SSE) on 8 of the

images. Since the controls now consist of both stratiform and convective pixels a stratiform/convective binary switch was used as the two external variables in Universal Kriging. This is indicated as a specialization of  $Q_c$  in Eq. (16b) and is given as Eq. (17):

$$Q_c = \begin{bmatrix} C_1 & S_1 \\ \vdots & \vdots \\ C_n & S_n \end{bmatrix} \quad (17)$$

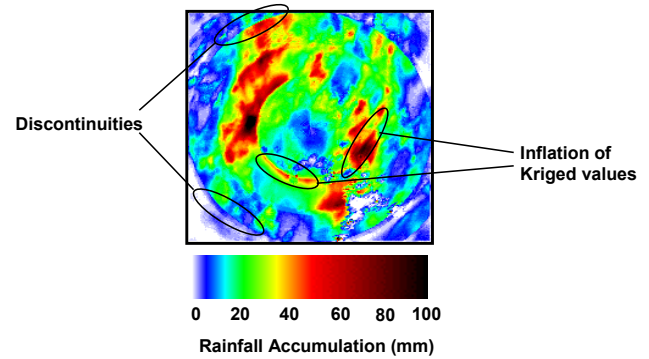
where  $C$  represents a convective and  $S$  a stratiform control; so that a convective control is assigned the values  $[C \ S]=[1 \ 0]$ ; and a stratiform control is assigned the values  $[C \ S]=[0 \ 1]$ .

A decision needs to be made as to which semivariogram is to be used in the coefficient matrix and which switch is to be set in  $Q_c$  (Eq. 17). This also affects the choice of  $q_t$  in the right hand side of Eq. (16a). It was decided to let this be determined by examining the controls; if the majority of the controls are convective then the target is assumed to be convective, if not then it is assumed to be stratiform. Including these external variables resulted in an improvement in the estimates. An example of this comparison is indicated in Fig. 13 where first Ordinary then Universal Kriging were used to estimate the mixed rainfall pixels on an instantaneous image (24 January 2002) from the Bethlehem weather radar; the number of pixels to be estimated were 2628. As indicated in Fig. 13, Universal Kriging provides an improved estimate over Ordinary Kriging.

#### 5.4 Cascade Kriging

When Kriging directly to ground level using the full set of CAPPI data in a stack, there are unexpected problems which occur. A 24-h accumulation of rainfall calculated from 5-min images Kriged from the CAPPI volume down to ground level is shown in Fig. 14. The data are from the Bethlehem weather radar (25 January 1996); the image is 200 km square. As can be seen in Fig. 14, there are serious discontinuities at the regions located directly under the edges of the CAPPIs; there is also an inflation of the reflectivity estimates which was found to be due to the numerical and geometric distribution of Kriging weights at the CAPPI edge locations.

A solution to the above problem is to use a Cascade Kriging approach, which exploits the fact that the 25 nearest neighbouring controls are confined to two levels above the target. The first step is to infill the missing data in the CAPPI at 16 km above ground level, pixel by pixel, using neighbourhood Kriging of the 25 control data vertically above and in the horizontal direction. Once all the unknown data on that CAPPI have been estimated and infilled, the CAPPI directly below is examined and any missing data here are estimated in the same manner; once again estimates from above (which now can include previously infilled data) and at the same level are used as control data. This is repeated until all the



**Fig. 14.** 24-h rainrate accumulation for the Bethlehem weather radar (25 January 1996). Image dimensions are 200 km square. Indicated on the image are the serious discontinuities associated with the CAPPI edges and also the inflation of the Kriged estimates at CAPPI edges.

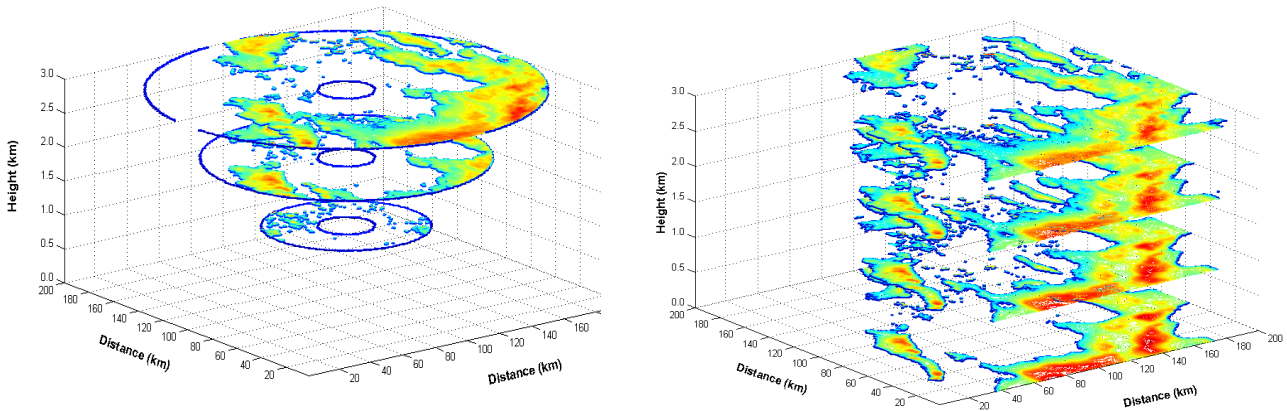
unknown data in each CAPPI are estimated. The final step involves an estimate of the rainfall at ground level.

Figure 15 gives a 3-D illustration of radar volume scan data before and after Cascade Kriging. The data are from the Bethlehem weather radar, 14 February 1996, where the image is 200 km square, and the vertical extent is 3 km. The levels above 3 km have been ignored for clarity, but the cascade was started at level 4. The image on the left hand side indicates the volume scan data before Cascade Kriging and the right hand image represents the same volume scan data after Cascade Kriging has been used to estimate as much unknown data as possible and provide an estimate at ground level.

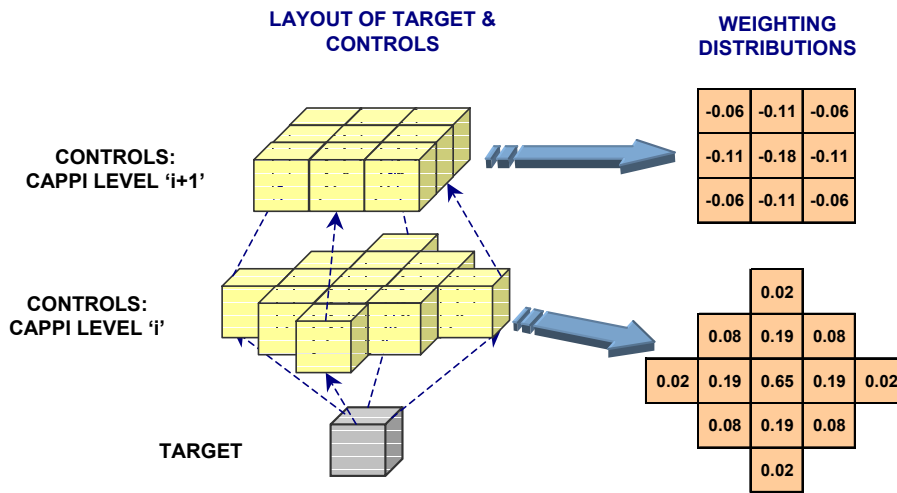
Cascade Kriging has the following influences on reflectivity data. There is increased smoothing of the reflectivity data as ground level is approached. There is also an increase in the reflectivity values as ground level is approached due to the pattern of Kriging Weights as illustrated in Fig. 16. This is because, in the case where the reflectivity values increase with proximity to ground level, the negative weights on the upper level and the positive weights on the lower level maintain the observed gradient, which results in an increase in the magnitude of the Kriged values; this tendency corresponds with the climatological profiles shown in Fig. 5. The next section details the validation testing procedure adopted to determine the efficiency and accuracy of the algorithm.

## 6 Testing and results

In order to determine the efficacy of the proposed technique in providing accurate rainfall accumulation values at ground level, a comparison was carried out between rain-gauge and radar accumulation values for selected rain events; care was taken to compute values which were comparable at ground level. The Liebenbergsvlei catchment described in



**Fig. 15.** 3-D Graphical illustration of radar volume scan data up to 3 km above ground level, before and after Cascade Kriging. Data are from the Bethlehem weather radar (14 February 1996). Data at levels 4 km and higher are not shown, but were used to successively fill the stack from above.



**Fig. 16.** Layout of Kriging weightings, computed with convective semivariogram parameters, for 22 controls (chosen for symmetry of illustration) centred directly above a target. The negative weights (occurring particularly where semivariograms model parameter  $\alpha > 2$ ) are situated on the upper CAPPI and the positive weights on the lower CAPPI and maintain a gradient in the field if there is one above the target.

the introduction and illustrated in Fig. 3 was used as the test area for the comparison between radar and raingauge accumulation estimates.

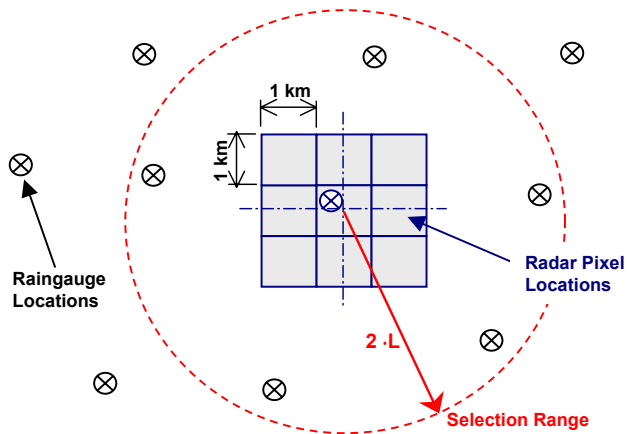
The radar rainfall estimates at ground level are estimated in 1 km by 1 km cells and the raingauge provides a point estimate at ground level. In order to provide an appropriate way of comparing the two estimates, the following approach was taken. The average value of the 9 radar pixels (a 3 km square area), with the centre square covering a gauge, was taken as the radar estimate of average rainfall at each gauge location. In computing the radar accumulation values one needs to take into account that the images are sampled at five-minute intervals, so a simple linear accumulation of the images may not provide an accurate accumulation, especially for fast moving rain events. The accumulations are therefore computed

by a Morphing Algorithm that takes into account the motion of the rainfield between instantaneous images (Sinclair and Pegram, 2003).

Block Kriging was used to determine the average rainfall over the same 9 km<sup>2</sup> area by using all raingauges within a range of 2 correlation lengths from the centre of the 9 km<sup>2</sup> area, at times corresponding to the radar estimates. The Block Kriging estimates are computed by Eq. (18):

$$\hat{z}_D = \frac{1}{D} \int_D Z(x) \cdot dx \tag{18}$$

where  $\hat{z}_D$  is the average value over a selected region and  $D$  defines the region (Bras and Rodriguez-Iturbe, 1985: 402–404). Due to the wide spacing of the raingauges compared



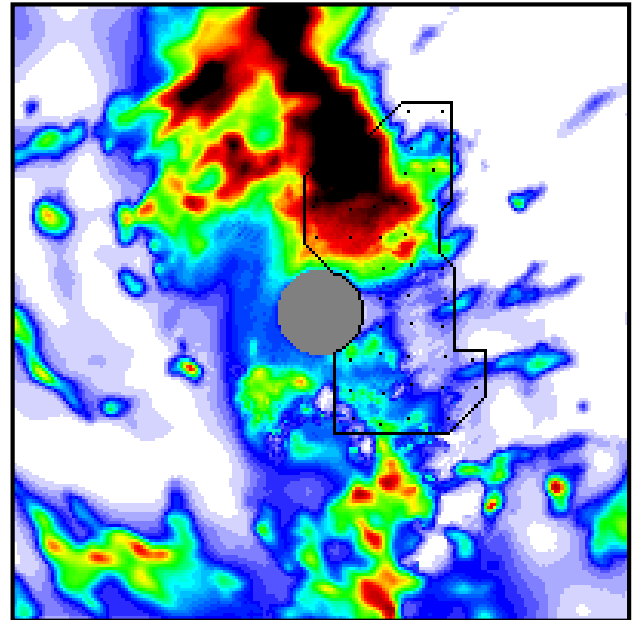
**Fig. 17.** Method of selection of raingauges to compute Block Kriging estimate over the 9 pixels centred at a raingauge location. Raingauges within a range of 2 correlation lengths are selected to compute the Block Kriging estimate over the 9 pixel region.

to the  $3 \times 3$  km area, the semivariogram parameters used were the same as those determined for the radar reflectivity data in the horizontal plane. However, a nugget effect is included, which for stratiform rain has a value of 0.10 and convective rainfall 0.03 (for a semivariogram type function) as was determined by Habib and Krajewski (2002). This allows for the random errors and uncertainties associated with raingauge estimates. Illustrated in Fig. 17 is the method of selection of the nearest raingauges to be used in the estimation of the average rainfall over the 9-pixel region, where the raingauges selected are within a range of 2 correlation lengths from the centre of the 9-pixel region.

The following two rain events were selected for comprehensive validation tests: (1) 24 January 1996 – a rain event which consisted of extremely high rainfall intensities over an approximately 12-h period, (2) 13 February 1996 – a rain event consisting of a combination of convective rain with periods of stratiform rain. Further testing on numerous other rain events was also conducted and is discussed in a later point in the Testing and Results section but not in as much detail as the two mentioned above.

### 6.1 Statistical comparison of raingauge and radar accumulations

In order to determine the quality of the radar estimates at ground level a statistical comparison of the raingauge and radar accumulations over different periods of 24 h, 12 h and 6 h was conducted. A scatter plot of the radar and raingauge accumulations was done and the correlation coefficient computed. The Kolmogorov-Smirnov (K-S) test was also used to determine if the cumulative distributions of the radar and raingauge accumulations were significantly different at a significance level of 5%, where the hypothesis being tested is



**Fig. 18.** 12-h rainfall accumulation from the Bethlehem radar on the 24 January 1996. An outline of the test catchment is indicated on the image as well as the raingauge locations. Dimensions of the image are 200 km square.

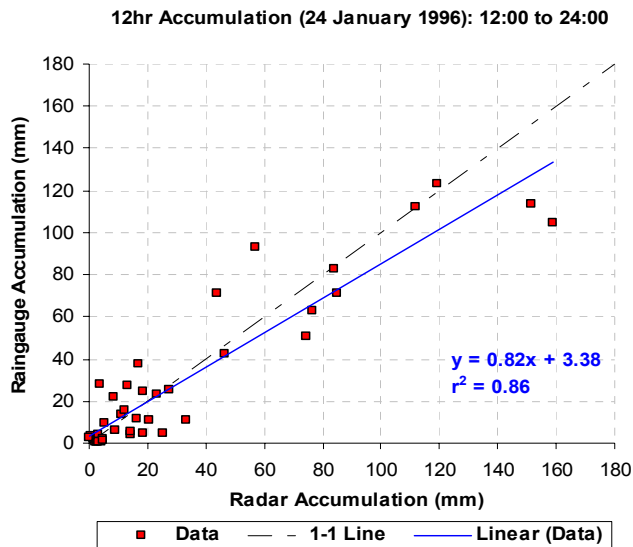
$H_0$ : Raingauge Distribution = Radar Distribution against  $H_1$ : Raingauge Distribution  $\neq$  Radar Distribution. The means and standard deviations ( $\sigma$ ) of the accumulated values were also tested at a significance level of 5%. In all the statistical tests undertaken a Box-Cox Transform was applied to the data in order to ensure the data sets approximated to a normal distribution.

## 6.2 Rainfall events

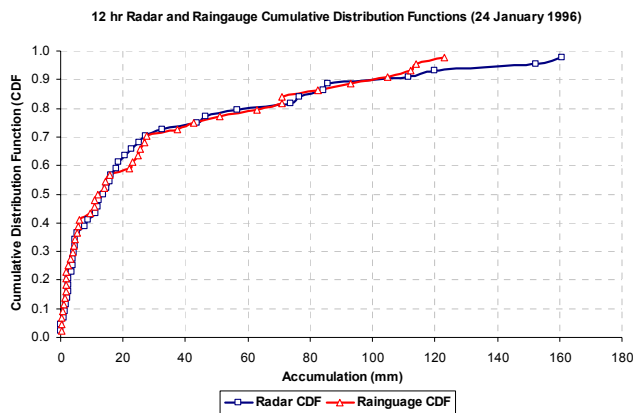
### 6.2.1 Rainfall event 1: 24 January 1996

Severe convective rainfall was recorded in the late afternoon and evening of the 24 January 1996 (Mather et al., 1997). The analysis for this event was conducted over a 12-h period from 12:00 to 24:00. For this rain event 29% of the images in the 12-h accumulation period were identified as containing bright band and corrected. Figure 18 illustrates a 12-h accumulation giving an indication of the extreme rainfall intensity during the 12-h period. Indicated on Fig. 18 is an outline of the catchment area as well as the raingauge locations. The region directly above the radar, a circle of 20 km radius, has also not been infilled due to poor data quality close to the radar during this time period; two raingauges within this region were excluded from the study.

A scatter plot of the raingauge and radar accumulations after a 12 h period at each of the 43 gauge sites computed by the appropriate Kriging method over each  $9 \text{ km}^2$  region, is shown in Fig. 19. There is a high correspondence for this



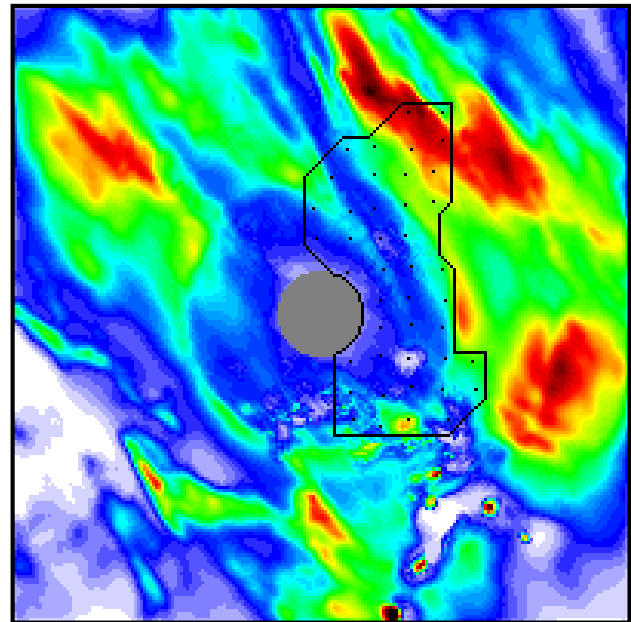
**Fig. 19.** Scatter plot of radar and raingauge accumulations for a 12-h accumulation period for the 24 January 1996 rain event. A strong correspondence exists between the radar and raingauge accumulation, especially for the high intensity rainfall.



**Fig. 20.** Cumulative Distribution Functions (CDFs) of the raingauge and radar accumulations for a 12-h period. Accumulations values are from the 24 January 1996 (12:00 to 24:00). Testing at a significance level of 5% indicate that the distributions of the accumulations for the radar and raingauge are not dissimilar.

rain event between radar and raingauge accumulations with an  $r^2$  value of 0.86 being returned. The highest pair of radar observations bias what is otherwise a reasonable fit, as can be seen in Figs. 19 and 20. This is probably due to under estimation of the highly localised convective rainfall by the gauges concerned (Wilson and Brandes, 1979).

Table 4 provides a summary of the statistical results returned for the raingauge and radar accumulation over a 12-h period and two 6-h periods. For the 6-h period (18:00 to 24:00) and the 12-h accumulation period there is a close correspondence between the radar and raingauge estimates.



**Fig. 21.** 24-h rainfall accumulation from the Bethlehem radar on the 13 February 1996. An outline of the test catchment is indicated on the image as well as the raingauge locations. Dimensions of the image are 200 km square.

With the exception of one standard deviation, the means and standard deviations are statistically similar with the K-S test also indicating that the distributions of the accumulations for the radar and raingauge are not dissimilar, except for the 6-h accumulation period from 12:00 to 18:00. The word “Accept” in Table 4 refers to the null hypothesis  $H_0$  being accepted and  $H_1$  rejected, and “Reject” refers to  $H_0$  being rejected and  $H_1$  accepted. Figure 20 shows an example of the cumulative distribution functions for 12-h raingauge and radar accumulations where the distributions for this example were considered to be not dissimilar.

#### 6.2.2 Rainfall event 2: 13 February 1996

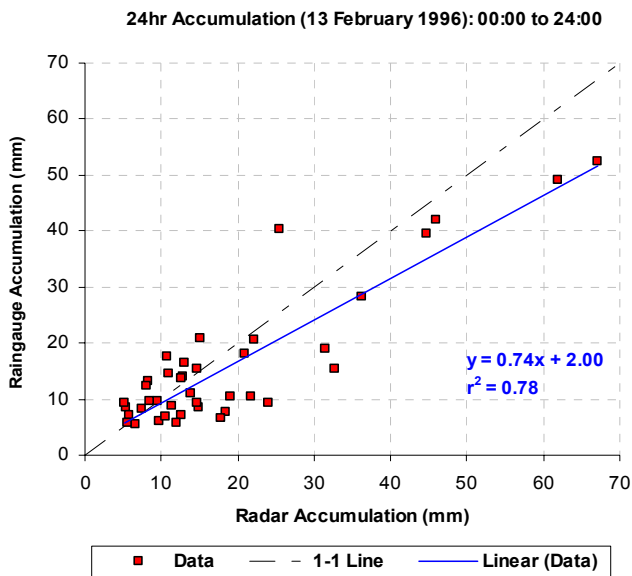
The rain event on the 13 February 1996 consisted of a combination of convective and stratiform rain. The rainfall initially consisted of light stratiform rainfall but towards the evening and afternoon this area also experienced convective rainfall. In the 24-h rain period 19% of the images were identified as containing bright band and corrected. The 24-h accumulation for the 13 February 1996 is depicted in Fig. 21 illustrating the combination of convective and stratiform rainfall that was recorded throughout the day. The exclusion of data directly above the radar applies in this instance as well. Two of the gauges were therefore excluded from the data analysis.

Figure 22 illustrates a scatter plot of the radar and raingauge areal average data for the 24-h accumulation period of low intensity, wide spread, mostly stratiform rain, corresponding to Fig. 21. The scatter plot indicates a strong



**Table 4.** Summary of statistical results for different accumulation periods for 24 January 1996 rain event for radar and raingauge data.

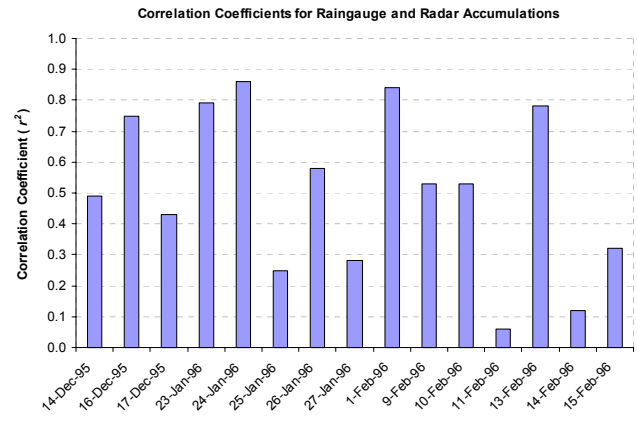
	6 h accumulation (12:00 to 18:00)		6 h accumulation (18:00 to 24:00)		12 h accumulation (12:00 to 24:00)	
	Mean	Stdev	Mean	Stdev	Mean	Stdev
Radar	4.6	5.1	27.2	40.8	30.3	41.3
Raingauge	4.0	7.1	25.4	38.3	29.4	36.8
Accept/Reject $H_0$	Accept	Accept	Accept	Accept	Accept	Accept
$r^2$	0.08		0.88		0.86	
K-S TEST	Reject		Accept		Accept	



**Fig. 22.** Scatter plot of radar and raingauge accumulations for a 24-h accumulation period for the 13 February 1996 rain event. A strong correspondence exists between the radar and raingauge accumulation, especially for the high intensity rainfall.

correspondence between raingauges and radar estimates with a correlation coefficient,  $r^2$ , of 0.78 being returned. The higher rainfall accumulation values show a strong correspondence, while the lower rainfall values exhibit a weaker relationship. Again, the raingauges underestimate the highest two values of the rainfall when compared to the radar, inducing a bias because of the strong influence of the two largest values.

Table 5 provides a summary of the statistical analysis of the comparisons between raingauges and radar accumulation for the full 24-h period and two 12-h periods. The mean and standard deviation for each of the accumulation periods closely corresponds and are statistically similar, except for the standard deviation of the 12-h accumulation from 12:00 to 24:00. The correlation coefficient value,  $r^2$ , is also high except for the accumulation period from 00:00 to 12:00



**Fig. 23.** Correlation coefficients ( $r^2$ ) computed for raingauge and radar accumulations over a 24-h period for 15 rain events consisting of different rainfall types.

which consisted of light stratiform rain. The cumulative distributions of accumulation values for each time period were also determined to be statistically similar via the K-S test.

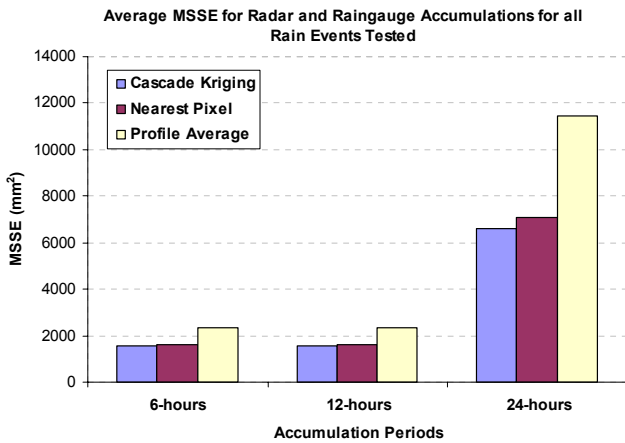
The statistics of the comparatively high rainfall accumulations summarised in Tables 4 and 5 indicate that the estimates of the rainfall at the ground from the radar images satisfactorily capture the amount and variability of the rainfall.

### 6.3 Additional testing

The Cascade Kriging algorithm was also tested on additional rain events that consisted of stratiform, convective and mixtures of convective and stratiform rainfall. The rain events were selected from the December 1995 and January/February 1996 periods all of which were taken from the Bethlehem weather radar and once again the Liebenbergsvlei raingauge network utilised. The rain events will not be discussed in as much detail as the two previously mentioned events however the correlation coefficient ( $r^2$ ) was computed and the mean, standard deviation and CDFs of the raingauge and radar accumulations were tested to see if they were statistically dissimilar. All tests were once again carried out at

**Table 5.** Summary of statistical results for different accumulation periods for 13 February 1996 rain event for radar and raingauge data.

	12 h accumulation (00:00 to 12:00)		12 h accumulation (12:00 to 24:00)		24 h accumulation (00:00 to 24:00)	
	Mean	Stdev	Mean	Stdev	Mean	Stdev
Radar	1.1	1.4	17.0	14.5	18.1	14.7
Raingauge	1.0	1.5	14.8	11.9	15.8	12.2
Accept/Reject $H_0$	Accept	Accept	Accept	Reject	Accept	Accept
$r^2$	0.14		0.76		0.78	
K-S TEST	Accept		Accept		Accept	



**Fig. 24.** Average MSSE for radar and raingauge accumulations for all rain events tested over 6, 12 and 24-h accumulation periods.

a significance level of 5%. In total 15 rain events are summarised in this sub-section.

The correlation coefficients were firstly computed for the raingauge and radar accumulations all of which were for a 24-h accumulation period. As mentioned in the rain events discussed above there was a good correspondence between raingauge and radar accumulations for convective rainfall events with a weaker correspondence for stratiform rainfall events. The correlation coefficients ( $r^2$ ) are plotted in Fig. 23.

For the 15 rain events tested 8 of the mean values for the raingauge and radar accumulations were considered to be similar with 7 dissimilar. For the computed standard deviations 8 were also considered to be similar with 7 dissimilar. The K-S test on the CDFs for the raingauge and radar accumulations returned that 6 of the distributions were statistically similar with 9 being considered dissimilar.

From the 15 rain events selected 8 consisted of predominantly stratiform rainfall and 7 with combinations of convective and stratiform rainfall. The average  $r^2$  value of the 8 stratiform rainfall events was 0.41 and the average  $r^2$  value from the 7 heavier rainfall events was 0.61. The major-

ity of the tests where the raingauge and radar accumulation statistics were considered to be dissimilar belonged to the predominantly stratiform rainfall events. The above results, and average  $r^2$  values computed, once again indicate that the Cascade Kriging extrapolation algorithm works well for the heavier rainfall events but returns less accurate results for lighter stratiform rainfall events.

#### 6.4 Comparison to alternative methodologies

A brief overview of how the Cascade Kriging algorithm compared to two simpler methodologies to obtain the rainrate at ground level will be reviewed.

- Nearest Pixel – The first method simply uses the lowest (altitude) pixels reflectivity in a vertical profile as the estimate at ground level.
- Average Profile – The second method is currently employed by SAWS in the Daily Rainfall Mapping Project over South Africa to provide an estimate the rainfall at ground level. The method calculates the average of all the non-zero reflectivity values contained in a vertical column above the pixel, the average is then utilised as the estimate of the rainfall at ground level (Visser, 2003).

The three methods: Cascade Kriging, Nearest Pixel and Profile Average were compared over 4 rain events (24 and 27 January 1996, 13 February 1996 and 17 December 1995) over the 6-h, 12-h and 24-h accumulation periods. For each of the methods employed to obtain an estimate of the rainfall at ground level over each of the accumulation periods the Mean Sum of Square of Errors (MSSE) was computed. The MSSE was determined by computing the SSE between the Block Kriged raingauge accumulations and radar accumulations and dividing by the number of raingauges utilised.

The Cascade Kriging algorithm returned preponderantly improved estimates over the other techniques in rain events that consisted of a mixture of stratiform and convective rainfall over the accumulation periods over a range of rainfall events. The average MSSE for each of the techniques over 6, 12 and 24-h accumulation periods was computed and the

results illustrated in Fig. 24. As indicated in Fig. 24 the Cascade Kriging algorithm returned superior estimates in terms of the MSSE over all the accumulation periods investigated; although it is marginally better than the Nearest Pixel method it is much better than the Profile Averaging.

## 7 Summary and conclusion

A technique has been presented that extrapolates radar volume scan data estimated at 1 km intervals above ground level to the earth's surface, and in the process estimates missing data in the CAPPI stack.

A rainfall classification was firstly done to separate the rainfall into stratiform and convective types. By separating the rainfall, climatological semivariograms for each rainfall type were defined, a procedure which ensures high computational efficiency with little loss in accuracy. The 2 km CAPPI is then examined to determine if a bright band is present and if so the pixels are adjusted on an individual basis. Both 3-D Universal and Ordinary Cascade Kriging were then used to infill missing data in the CAPPI stack and finally provide a rainfall estimate at ground level. In the Kriging computations care was taken to ensure computational efficiency and stability.

The effectiveness of the above algorithm was tested on two different rainfall events exhibiting distinctly different types of rainfall by comparing raingauge and radar accumulation estimates at ground level. The technique demonstrated that both high and moderate intensity rainfall exhibited similar estimates to raingauges whereas stratiform rainfall showed a weaker correspondence to the raingauge estimates; the research is ongoing. The Cascade Kriging technique is undergoing further testing currently at the Innovation and Research Division of SAWS to determine if it is operationally sound to implement.

*Acknowledgements.* This work is currently supported financially by the South African Water Research Commission (WRC) and the National Research Foundation (NRF). Thanks go to S. Sinclair for providing the computer code for the Morphing Algorithm to compute the radar accumulation values. Thanks must also go to meteorologists D. Terblanche and P. Visser of SAWS at Bethlehem, South Africa for invaluable advice regarding the rainfall classification and bright band correction algorithms.

Edited by: P. P. Alberoni

Reviewed by: two referees

## References

- Bras, R. L. and Rodriguez-Iturbe, I.: Random Functions and Hydrology, Addison-Wesley Publishing Company, Ontario, 1985.
- Ciach, G. J.: Local Random Errors in Tipping-Bucket Rain Gauge Measurements, *J. Atmos. Oceanic Technol., Notes and Correspondence*, 20, 752–759, 2002.
- Chiles, J. P. and Delfiner, P.: Geostatistics – Modeling Spatial Uncertainty, Wiley, New York, USA, 1999.
- Cressie, N.: Statistics for Spatial Data, Wiley-Interscience Publication, New York, USA, 1993.
- Cressie, N. and Hawkins, D. M.: Robust Estimation of the Variogram, 1., *J. Int. Assoc. Mathematical Geology*, 12, 115–125, 1980.
- Creutin, J. D. and Obled, C.: Objective Analysis and Mapping Techniques for Rainfall Fields: An Objective Comparison, *Water Resour. Res.*, 18(2), 413–431, 1982.
- Franco, M., Sempere-Torres, D., Sanchez-Diezma, R., and Andrieu, H.: A methodology to identify the vertical profile of reflectivity from radar scans and to estimate the rainrate at ground at different distances, *Proc. ERAD (2002)*, 299–304, 2002.
- Gordon, A. D.: Classification: Methods for the exploratory analysis of multivariate data, Chapman and Hall, New York, USA, 1981.
- Habib, E. and Krawjewski, W. F.: Uncertainty Analysis of the TRMM Ground-Validation Radar-Rainfall Products: Application to the TEFLUN-B Field Campaign, *J. Appl. Meteorol.*, 41, 558–572, 2002.
- Hengl, T., Geuvelink, G. B. M., and Stein, A.: Comparison of Kriging with external drift and regression-kriging, Technical note, ITC, available online at: [http://www.itc.nl/library/Academic\\_output/](http://www.itc.nl/library/Academic_output/), 2003.
- Jordan, P., Seed, A., and Austin, G.: Sampling errors in radar estimates of rainfall, *J. Geophys. Res.*, 105, D2, 2247–2257, 2000.
- Journel, A. G. and Huijbrechts, C. J.: Mining Geostatistics, Academic Press, London, UK, 1978.
- Lack, S. A. and Fox, N. I.: Errors in Surface Rainfall Rates Retrieved from Radar due to Wind Drift, Sixth International Symposium on Hydrological Applications of Weather Radar, Melbourne, Australia, 2004.
- Lebel, T., Bastin, G., Obled, C., and Creutin, J. D.: On the Accuracy of Areal Rainfall Estimation: A Case Study, *Water Resour. Res.*, 23, 11, 2123–2134, 1987.
- Marshall, J. S. and Palmer, W. M.: The distribution of raindrops with size, *J. Meteorol.*, 5, 165–166, 1948.
- Mather, G. K., Terblanche, D. E., and Steffens, F. E.: National Precipitation Research Programme: Final Report for the Period 1993–1996, WRC report No 726/1/97, 1997.
- Matheron, G.: *Traite de Geostatistique Appliquee*, Tome I., *Memoires du Bureau de Recherches Geologiques et Minières*, 14, Editions Technip, Paris, 1962.
- Mittermaier, M. P.: Investigating the Characteristics of the Radar Reflectivity Profile, MSc Eng dissertation, Civil Engineering, University of KwaZulu-Natal, Durban, Howard College, 1999.
- Mittermaier, M. P.: Investigating synergies between weather radar data and mesoscale model output, PhD Thesis, Department of Meteorology, University of Reading, 2003.
- Sabyasachi, B., Gunst, R. F., Guertal, E. A., and Hartfield, M. I.: The Effects of Influential Observations on Sample Semivariograms, *J. agriculture, biological, and environmental statistics*, 2, 490–512, 1997.
- Sanchez-Diezma, R., Zawadzki, I., and Sempere-Torres, D.: Identification of the bright band through the analysis of volumetric radar data, *J. Geophys. Res.*, 105(D2), 2225–2236, 2000.
- Seed, A. W. and Pegram, G. G. S.: Using Kriging to infill gaps in radar data due to ground clutter in real time, *Proc. Fifth Int. Symposium on Hydrological Applications of Weather Radar –*

- Radar Hydrology, Kyoto, Japan, 73–78, 2001.
- Sempere-Torres, D., Sanchez-Diezma, R., Zawadzki, I., and Creutin, J. D.: Identification of Stratiform and Convective Areas Using Radar Data with Application to the Improvement of DSD Analysis and Z-R Relationships, *Phys. Chem. Earth*, 25, 985–990, 2000.
- Sinclair, D. S. and Pegram, G. G. S.: Combining traditional and remote sensing techniques of rainfall measurement as a tool for Hydrology, Agriculture and Water Resources Management, Proc. 11th SA National Hydrology Symposium, Port Elizabeth, South Africa, 2003.
- Steiner, M., Houze, R. A., and Yuter, S. E.: Climatological Characterization of Three-Dimensional Storm Structure from Operational Radar and Rain Gauge Data, *J. Appl. Meteorol.*, 34, 1978–2007, 1995.
- Terblanche, D. E., Pegram, G. G. S., and Mittermaier, M. P.: The development of weather radar as a research and operational tool for hydrology in South Africa, *J. Hydrol.*, 241, 3–25, 2001.
- Visser P.: Spatial interpolation and Mapping of rainfall: 2. Radar and Satellite products, Draft final report (K5/1152) to the Water Research Commission for the period March 2000 to March 2003, Water Research Commission, Pretoria, 19–29, 2003.
- Wesson, S. M. and Pegram, G. G. S.: Radar rainfall image repair techniques, *Hydrol. Earth Syst. Sci.*, 8, 220–234, 2004.
- Wilkinson, J. H.: *The Algebraic Eigenvalue Problem*, Oxford Science Publication, 1988.
- Wilson, J. W. and Brandes, E. A.: Radar measurement of rainfall – A summary, *Bull. Amer. Meteorol. Soc.*, 60(9), 1048–1058, 1979.