



## Uncertainties Associated with Arithmetic Map Operations in GIS

JORGE K. YAMAMOTO<sup>1</sup>, ANTÔNIO T. KIKUDA<sup>1</sup>, GUILHERME J. RAMPAZZO<sup>1</sup> and CLAUDIO B.B. LEITE<sup>2</sup>

<sup>1</sup>Universidade de São Paulo, Instituto de Geociências, Geologia Sedimentar e Ambiental, Rua do Lago, 563, Cidade Universitária, 05508-900 São Paulo, SP, Brazil

<sup>2</sup>Universidade Federal de São Paulo, Departamento de Ciências Exatas e da Terra, Rua São Nicolau, 210, 09913-030 Diadema, SP, Brazil

*Manuscript received on August 3, 2017; accepted for publication on October 9, 2017*

### ABSTRACT

Arithmetic map operations are very common procedures used in GIS to combine raster maps resulting in a new and improved raster map. It is essential that this new map be accompanied by an assessment of uncertainty. This paper shows how we can calculate the uncertainty of the resulting map after performing some arithmetic operation. Actually, the propagation of uncertainty depends on a reliable measurement of the local accuracy and local covariance, as well. In this sense, the use of the interpolation variance is proposed because it takes into account both data configuration and data values. Taylor series expansion is used to derive the mean and variance of the function defined by an arithmetic operation. We show exact results for means and variances for arithmetic operations involving addition, subtraction and multiplication and that it is possible to get approximate mean and variance for the quotient of raster maps.

**Key words:** GIS, Interpolation variance, Map Algebra, Propagation of Uncertainty, Taylor series.

### INTRODUCTION

A map resulting from interpolation of field data must have some assessment of uncertainty (e.g. Heuvelink et al. 1989, Crossetto et al. 2000, Atkinson and Foody 2002). When estimates are derived from field data, they have associated uncertainties caused by spatial variation of continuous variables, interactions among them and the effect of neighbor data (Wang et al. 2005). Generally, kriging is used to predict values of the variable of interest at unsampled locations, because it provides an assessment of uncertainty as given by the kriging variance. However, the kriging variance does not depend on real local data values (Atkinson and Foody 2002) and therefore this is not a reliable measure of local accuracy. According to Wang et al. (2005), local estimates are strongly associated with neighbor data. Actually, the kriging variance is just a measure of the ranking of data configurations given by the variogram model (Journel and Rossi 1989). A measure of local

---

Correspondence to: Guilherme José Rampazzo  
E-mail: [grampazzo@usp.br](mailto:grampazzo@usp.br)

\* Contribution to the centenary of the Brazilian Academy of Sciences.



accuracy should take into account the data value. In this sense, Yamamoto (2000) proposed interpolation variance as a measure of the reliability of kriging estimates taking into account both data configuration and data values.

Moreover, interpolation variance can also be used for quantifying uncertainty associated with interpolation of categorical variable (Yamamoto et al. 2012). Therefore, interpolation variance is valid for both continuous and discrete variables. Actually, the interpolation variance can be used with any interpolation method based on weighted average formula (for instance, inverse of distance weighting, multiquadric equations, etc.). Yamamoto et al. (2012) used interpolation variance as a measure of uncertainty associated with multiquadric interpolation. Because interpolation is based on limited information given by a sample, the resulting raster map is always subject to uncertainty. Usually, the main source of uncertainty is related to a lack of knowledge, but there are other factors affecting the final uncertainty in a raster map. Burrough (1986) lists 14 factors affecting the uncertainty. These factors were classified into three groups: I) obvious sources of error; II) errors resulting from natural variation of original measurements; III) errors arising through processing (Burrough 1986). Wellmann et al. (2010) classified sources of uncertainty affecting geological data into three categories: I) imprecision and measurement error; II) stochastic nature of the geological variable; III) imprecise knowledge. Considering the field data with negligible uncertainty, the measured variance at unsampled location takes into account uncertainty due to lack of knowledge because of insufficient sampling (three factors of Burrough (1986), affect the interpolation variance: density of observations; natural variation and interpolation).

Considering that we have an interpolated raster map and its associated uncertainty, we can perform arithmetic operations with raster maps. Arithmetic map operations are very common procedures used in geographic information system. In general, two raster maps are combined using arithmetic operators (addition, subtraction, multiplication and division) to derive a new raster map. Let  $X$  and  $Y$  be interpolated raster maps from the same data set and let us perform some arithmetic operation between them. For each variable we know not only the interpolated value but also the uncertainty given by the interpolation variance. Thus, we have to calculate the resulting raster map after an arithmetic operation and the uncertainty as well. Uncertainty is magnified when cartographic overlay operations involve more than two steps (Burrough 1986). We can define a function of variables  $X$  and  $Y$ :  $f(x, y)$  that can be expanded in a Taylor series about the mean values of  $X$  and  $Y$ . From this expansion we can apply the mathematical expectation operator to find the mean value of the function and the definition of the variance as a measurement of the uncertainty. Addition and subtraction are very simple arithmetic operators, but multiplication and division are much more complex operations and they require correcting the combined estimate and the calculation of the variance is even more complicated.

#### ESTIMATES AND VARIANCES OF ARITHMETICALLY COMBINED VARIABLES

Given a random variable  $X$ , the mathematical expectation is:

$$E[X] = \frac{1}{n} \sum_{i=1}^n X_i \quad (1)$$

The variance is a measure of the uncertainty associated with the mathematical expectation:

$$\text{Var}[X] = \frac{1}{n} \sum_{i=1}^n (X_i - E[X])^2 \quad (2)$$

which can be written as:

$$\text{Var}[X] = E[X - E[X]]^2 = E[X^2] - (E[X])^2 \quad (3)$$

Given two random variables (X and Y), we can compute the covariance as:

$$\text{Cov}(X, Y) = E[XY] - E[X]E[Y] \quad (4)$$

The correlation coefficient as a measure of the mutual relationship between two random variables is calculated as:

$$\rho_{X,Y} = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}[X] \text{Var}[Y]}} = \frac{\text{Cov}(X, Y)}{S_X S_Y} \quad (5)$$

Where  $S_x$  and  $S_y$  are standard deviations.

Let us call  $f(x, y)$  a function resulting from arithmetically combined random variables X and Y. We are interested in the mean value of the function  $f(x, y)$  around the mean values of random variables X and Y. To find the mean and variance of this function we use Taylor expansion around  $\theta$ , which is the point near the mean values of X and Y:  $\theta = (\mu_x, \mu_y)$ . For the following development let us consider a simplified notation in which:  $\mu_x = E[X]$ ;  $\mu_y = E[Y]$ ;  $\sigma_x^2 = \text{Var}[X]$ ;  $\sigma_y^2 = \text{Var}[Y]$  and  $\sigma_{xy} = \text{Cov}(X, Y)$ . Notice that for propagation of uncertainty we need to know the means, variances, the covariance, and the type of operation (Heuvelink et al. 1989).

The second-order Taylor expansion is (Weir and Hass 2014):

$$\begin{aligned} f(x, y) = & f(\theta) + f_x(\theta)(x - \mu_x) + f_y(\theta)(y - \mu_y) \\ & + \frac{1}{2}f_{xx}(\theta)(x - \mu_x)^2 + f_{xy}(\theta)(x - \mu_x)(y - \mu_y) + \frac{1}{2}f_{yy}(\theta)(y - \mu_y)^2 + \text{remainder} \end{aligned} \quad (6)$$

Applying the expectation operator in (5) and neglecting the remainder, we have:

$$E[f(x, y)] = f(\theta) + \frac{1}{2}f_{xx}(\theta)\sigma_x^2 + f_{xy}(\theta)\sigma_{xy} + \frac{1}{2}f_{yy}(\theta)\sigma_y^2 \quad (7)$$

This is the general expression for computing the mathematical expectation of the function  $f(x, y)$  about  $\theta = (\mu_x, \mu_y)$ .

The variance of the function  $f(x, y)$  around  $\theta = (\mu_x, \mu_y)$  can be computed as:

$$\text{Var}(f(x, y)) = E[(f(x, y) - E(f(x, y)))^2] \quad (8)$$

Replacing definitions (6) and (7) in (3), we have:

$$\begin{aligned} \text{Var}[f(x, y)] = & E[(f_x(\theta)(x - \mu_x) + f_y(\theta)(y - \mu_y) + \frac{1}{2}f_{xx}(\theta)(x - \mu_x)^2 + f_{xy}(\theta)(x - \mu_x)(y - \mu_y) \\ & + \frac{1}{2}f_{yy}(\theta)(y - \mu_y)^2 - \frac{1}{2}f_{xx}(\theta)\sigma_x^2 - f_{xy}(\theta)\sigma_{xy} - \frac{1}{2}f_{yy}(\theta)\sigma_y^2)^2] \end{aligned} \quad (9)$$

**TABLE I**  
**Mean and variance for the function  $f(x, y)$  around  $\theta = (\mu_x, \mu_y)$ .**

$f(x, y)$	Derivatives	Mean	Variance		
$X \pm Y$	$f_x(x, y) = 1$	$E[f(x, y)] = \mu_x \pm \mu_y$ (10)	$Var[f(x, y)] = \sigma_x^2 + \sigma_y^2 \pm 2\sigma_{xy}$ (11)		
	$f_y(x, y) = \pm 1$				
	$f_{xx}(x, y) = 0$				
	$f_{yy}(x, y) = 0$				
	$f_{xy}(x, y) = 0$				
$XY$	$f_x(x, y) = y$	$E[f(x, y)] = \mu_x \mu_y + \sigma_{xy}$ (12)	$Var[f(x, y)] = \mu_y^2 \sigma_x^2 + \mu_x^2 \sigma_y^2 + E[(x - \mu_x)^2 (y - \mu_y)^2] - \sigma_{xy}^2 + 2\mu_x \mu_y \sigma_{xy} + 2\mu_y E[(x - \mu_x)^2 (y - \mu_y)] + 2\mu_y E[(x - \mu_x)(y - \mu_y)^2]$ (13)		
	$f_y(x, y) = x$				
	$f_{xx}(x, y) = 0$				
	$f_{yy}(x, y) = 0$				
	$f_{xy}(x, y) = 1$				
$X/Y$	$f_x(x, y) = \frac{1}{y}$	$E[f(x, y)] \approx \frac{\mu_x}{\mu_y} - \frac{\sigma_{xy}}{\mu_y^2} + \frac{\mu_x \sigma_y^2}{\mu_y^3}$ (14)	$Var[f(x, y)] \approx \frac{\sigma_x^2}{\mu_y^2} + \frac{\mu_x^2}{\mu_y^4} \sigma_y^2 - 2\frac{\mu_x}{\mu_y^3} \sigma_{xy}$ (16)		
	$f_y(x, y) = \frac{-x}{y^2}$				
	$f_{xx}(x, y) = 0$				
	$f_{yy}(x, y) = \frac{2x}{y^3}$				
	$f_{xy}(x, y) = \frac{-1}{y^2}$			$E[f(x, y)] \approx \frac{\mu_x}{\mu_y} - \frac{\sigma_{xy}}{\mu_y^2} + \frac{\mu_x \sigma_y^2}{\mu_y^3} + \frac{E[(x - \mu_x)(y - \mu_y)^2]}{\mu_y^3} + \frac{\mu_x E[(y - \mu_y)^3]}{\mu_y^4}$ (15)	$Var[f(x, y)] \approx \frac{\sigma_x^2}{\mu_y^2} + \frac{\mu_x^2}{\mu_y^4} \sigma_y^2 - 2\frac{\mu_x}{\mu_y^3} \sigma_{xy} + \frac{1}{\mu_y^4} E[(\sigma_{xy} - (x - \mu_x)(y - \mu_y))^2] + \frac{\mu_x^2}{\mu_y^6} E[((y - \mu_y)^2 - \sigma_y^2)^2] + \frac{2}{\mu_y^5} E[(x - \mu_x)[\sigma_{xy} - (x - \mu_x)(y - \mu_y)]] + 2\frac{\mu_x}{\mu_y^4} E[(x - \mu_x)[(y - \mu_y)^2 - \sigma_y^2]] - 2\frac{\mu_x}{\mu_y^4} E[(y - \mu_y)[\sigma_{xy} - (x - \mu_x)(y - \mu_y)]] - 2\frac{\mu_x^2}{\mu_y^5} E[(y - \mu_y)[(y - \mu_y)^2 - \sigma_y^2]] + 2\frac{\mu_x}{\mu_y^5} E[[\sigma_{xy} - (x - \mu_x)(y - \mu_y)][(y - \mu_y)^2 - \sigma_y^2]]$ (17)
	$f_{xxx}(x, y) = 0$				
	$f_{xxy}(x, y) = 0$				
	$f_{xyy}(x, y) = \frac{2}{y^3}$				
$f_{yyy}(x, y) = \frac{-6x}{y^4}$					

This expression cannot be simplified because it depends on partial derivatives of the function  $f(x, y)$ .

The mean and variance of the function  $f(x, y)$  are presented in Table I. All formulas result from second-order Taylor expansion (except for equations (13) and (14)) and are equal to those derived by Mood et al. (1974). Equation (13) results from third-order Taylor expansion. This equation should provide a better approach for the mean of quotient of random variables. The variance after equation (14) uses only first-order terms (Mood et al. 1974).

It can be noted that for the quotient of X and Y, the mean and variance are calculated by approximate formulas. For a better approximation of the variance of the quotient of X and Y, we considered second-order terms according to Maskey and Guinot (2003). The variance of the quotient after equation (15) is better than the variance computed after equation (14). Indeed, the algebraic complexity increased very much with the inclusion of second-order terms (Maskey and Guinot 2003). The mean and variance for the quotient of X and Y will always result in approximate formulas because the function  $f(x, y) = X/Y$  is infinitely differentiable with respect to Y. An exact formula for the calculation of the variance of the quotient of X and Y was proposed by Frishman (1971). However, this alternative uses the variance of the product  $X * 1/Y$ , which is implied in estimating  $1/Y$  instead of  $Y$ . Details of the mathematical development of Taylor expansion to find the mean and variance of the function  $f(x, y)$  can be found in the Appendix. This development was based entirely on statistics (Mood et al. 1974) and calculus (Weir and Hass 2014).

Heuvelink et al. (1989) proposed a general formula based on second order Taylor expansion for n variables. The mean of the function  $f(M)$  around  $M = \{\mu_1, \mu_2, \dots, \mu_n\}$  is computed as (Heuvelink et al. 1989):

$$\begin{aligned}
 E(f(M)) &= E[f(M)] + E\left[\sum_{i=1}^n [(x_i - \mu_i) f_i(M)]\right] + \frac{1}{2} E\left[\sum_{i=1}^n \sum_{j=1}^n [(x_i - \mu_i)(x_j - \mu_j) f_{ij}(M)]\right] \\
 &= f(M) + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n [\rho_{ij} \sigma_i \sigma_j f_{ij}(M)]
 \end{aligned}
 \tag{19}$$

Where  $\rho_{ij}$  is the correlation coefficient between variables  $i$  and  $j$ ;  $\sigma_i$  is the standard deviation for variable  $i$  and  $\sigma_j$  is the standard deviation for variable  $j$ .

The general formula for variance of the function  $f(M)$  is (Heuvelink et al. 1989):

$$\text{Var}(\hat{x}) = E\left\{ \left[ \sum_{k=1}^n [(x_k - \mu_k) f_k(M)] + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n [(x_i - \mu_i)(x_j - \mu_j) - \rho_{ij} \sigma_i \sigma_j f_{ij}(M)] \right]^2 \right\}
 \tag{20}$$

$$\begin{aligned}
 \text{Var}(\hat{x}) &= \sum_{k=1}^n \sum_{l=1}^n \rho_{kl} \sigma_k \sigma_l f_k(M) f_l(M) \\
 &+? \sum_{k=1}^n \sum_{i=1}^n \sum_{j=1}^n \{E[(x_k - \mu_k)((x_i - \mu_i)(x_j - \mu_j) - \rho_{kl} \sigma_k \sigma_l)]\} (f_k(M)) * f_{ij}(M) \} \\
 &+? \frac{1}{4} \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n \sum_{l=1}^n \{E[(x_i - \mu_i)(x_j - \mu_j)(x_k - \mu_k)(x_l - \mu_l) - \rho_{ij} \sigma_i \sigma_j \rho_{kl} \sigma_k \sigma_l] f_{ij}(M) f_{kl}(M)\}
 \end{aligned}
 \tag{21}$$

We proved that all formulas presented in Table I (except equations (12) and (14)) are equal to those developed from Heuvelink’s general formulas for mean and variance (Appendix).

**LOCAL ESTIMATES AND UNCERTAINTIES USING MULTIQUADRIC EQUATIONS**

This paper concerns building raster maps from field data, which means interpolation of regularly spaced cells based on neighbor data. Multiquadric equations (Hardy 1971) were chosen for interpolation of the variable at unsampled locations. Usually, several variables are measured in each sample location. For instance, in geochemical exploration, major and trace elements are analyzed simultaneously for each sample. Once we have built raster maps with these variables, we can combine variables to get another raster map. This is a common procedure in GIS. We present a method that allows combination of random variables and also uncertainty quantification based on equations summarized in Table I. It is important to emphasize that it depends on a reliable measure of uncertainty. In this sense, we propose the use of the interpolation variance (Yamamoto 2000) as an approach for uncertainty quantification. Interpolation variance can be used with any estimate based on weighted average formula.

For each cell location within the raster map, we have to find the nearest neighbor data. The dual form of multiquadric interpolation can be written as (Yamamoto 2002):

$$Z_{MQ}^*(x_o) = \sum_{i=1}^n w_i Z(x_i) \quad (22)$$

The weights  $\{w_i, i = 1, \dots, n\}$  come from the solution of a system of linear equations (Yamamoto 2002):

$$\sum_{j=1}^n w_j \theta(x_i - x_j) + \mu = \theta(x_i - x_o) \text{ for } i = 1 \quad \sum_{j=1}^n w_j = 1 \quad (23)$$

Where  $\mu$  is the Lagrange multiplier and  $\theta(x) = \sqrt{|x|^2 + C}$  is the multiquadric kernel as proposed by Hardy (1971).

Because equation (17) is a weighted average formula, we can use the interpolation variance for quantifying the uncertainty associated with multiquadric interpolation. The interpolation variance is simply (Yamamoto 2000):

$$S_o^2(x_o) = \sum_{i=1}^n w_i [Z(x_i) - Z_{MQ}^*(x_o)]^2 \quad (24)$$

Raster map operations consist of combining two variables at once. Let us call  $Z_1(x)$  and  $Z_2(x)$  the two variables to be arithmetically combined. Both variables are measured on same data location and the data set has the same number of data points. This is the essential condition to perform arithmetical operation between variables  $Z_1(x)$  and  $Z_2(x)$ . Since we are using multiquadric equations with the same multiquadric kernel, then the weights are identical:  $\{w_i^1 = w_i^2, i = 1, \dots, n\}$ . In light of this fact we will consider hereafter the set of weights  $\{w_i, i = 1, \dots, n\}$  for both variables.

Multiquadric interpolations for variables  $Z_1(x)$  and  $Z_2(x)$  are given as:

$$Z_{MQ_1}^*(x_o) = \sum_{i=1}^n w_i Z_1(x_i) \quad (25)$$

$$Z_{MQ_2}^*(x_o) = \sum_{i=1}^n w_i Z_2(x_i) \quad (26)$$

The term  $E[Z_1(x)Z_2(x)]$  as a local statistic can be calculated as:

$$E[Z_1(x)Z_2(x)] = \sum_{i=1}^n (w_i Z_1(x_i)Z_2(x_i)) \quad (27)$$

Thus, the local covariance can be computed as:

$$Cov(Z_1(x), Z_2(x)) = \sum_{i=1}^n w_i Z_1(x_i)Z_2(x_i) - \left( \sum_{i=1}^n w_i Z_1(x_i) \right) \left( \sum_{i=1}^n w_i Z_2(x_i) \right) \quad (28)$$

See that equation (27) allows computing cell-by-cell covariance. This statistic is fundamental for propagation of uncertainty. The interpolation covariance in (27) is calculated directly from neighbor data using the same weights for multiquadric interpolation. Therefore, this approach is different from the spatial covariance derived from the variogram model. Actually, for  $Z_1(x) = Z_2(x)$ , this formula is equivalent to the interpolation variance.

Interpolation variances are computed as:

$$S_{o_1}^2 = \sum_{i=1}^n w_i [Z_1(x_i) - Z_{MQ}^*(x_o)]^2 \tag{29}$$

$$S_{o_2}^2 = \sum_{i=1}^n w_i [Z_2(x_i) - Z_{MQ}^*(x_o)]^2 \tag{30}$$

Means and variances in Table I concern global statistics. But, raster maps imply local statistics and the purpose is to show that local statistics can be combined to compute mean and variances after mathematical operations between two random variables.

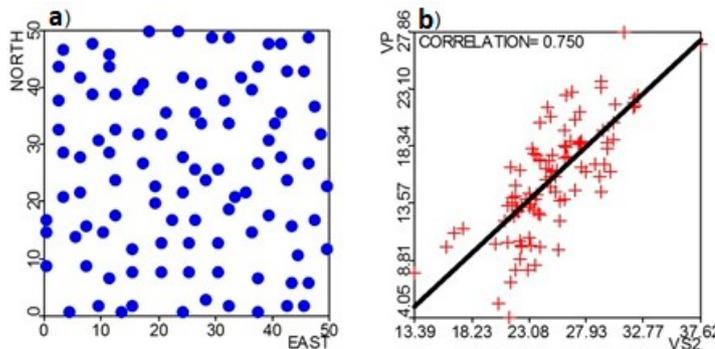
Keeping the same notation as used in Table I, we consider the following equivalence:  $\mu_x = Z_{MQ_1}^*(x)$ ;  $\mu_y = Z_{MQ_2}^*(x)$ ;  $\sigma_x^2 = S_{o_1}^2(x_o)$ ;  $\sigma_y^2 = S_{o_2}^2(x_o)$  and  $\sigma_{xy} = Cov(Z_1(x), Z_2(x))$ .

In addition, high-order moments in Table I are calculated as:

$$E[.] = \sum_{i=1}^n w_i(.)$$

**MATERIALS AND METHODS**

For this study we consider a stratified random sample drawn from an exhaustive data set composed of 2500 data points arranged in an array of 50 rows by 50 columns. This sample, with 100 data points (4% of the exhaustive) presents two measured variables (VP and VS2) in each location (Figure 1a). These variables present a positive correlation (Figure 1b). In addition to performing arithmetic operations between VP and VS2, we have to prove that mean and variances of combined variables are valid. Thus, we generated four new variables: VP+VS2; VP-VS2; VPxVS2 and VP/VS2. Sample statistics for all variables are presented in Table II.



**Figure 1** - Location map of sample data (a) and VP X VS2 scattergram (b) (see the colors in the online version).

VP and VS2 will be used for building raster maps, from which we want to obtain new raster maps by combining them arithmetically. Mean and variance for new random variables are listed in Table III. Comparing these results with directly computed statistics in Table II, we verify that except for the quotient of random variables, all others are exactly the same.

The sample data with six variables are the materials of this study. Original variables VP and VS2 will be used to build raster maps based on multiquadric interpolation. Equations (24) and (28) will be used

**TABLE II**  
**Sample statistics for variables VP, VS2, VP+VS2, VP-VS2, VPxVS2 and VP/VS2.**

Statistics	VP	VS2	VP+VS2	VP-VS2	VPxVS2	VP/VS2
Nb. of data	100	100	100	100	100	100
Mean	15.568	25.051	40.619	-9.483	403.047	0.615
Variance	21.130	14.236	61.486	9.245	28571.847	0.017
Coef. Var.	0.295	0.151	0.193	-0.321	0.419	0.215
Max.	27.863	37.620	64.377	-3.087	1006.605	0.893
UQ	18.421	27.203	45.624	-7.647	499.767	0.694
Median	15.774	24.666	40.951	-9.326	394.319	0.625
LQ	12.757	22.391	35.219	-11.679	288.545	0.543
Min.	4.049	13.386	21.134	-17.389	86.794	0.189

UQ=upper quartile and LQ=lower quartile.

to compute cells of the VP raster maps and equations (25) and (29) for VS2 raster maps. The covariance between VP and VS2 will be calculated after equation (27). These statistics will be combined according to formulas in Table I to derive means and uncertainties of resulting arithmetically combined raster maps. Now, we can compare resulting raster maps with those computed directly from new variables based on equations (16) and (17). Next, these equations ((21) and (23)) will be checked against equations (10) and (11) for variables  $VP \pm VS2$ , equations (12) and (13) for variable  $VPxVS2$  and equations (14), (15), (16) and (17) for variable  $VP/VS2$ .

## RESULTS AND DISCUSSION

Raster maps for variables VP and VS2 are presented in Figure 2. For the purpose of displaying interpolation variances and covariance, we considered a maximum equal to 95% of their distributions. Now, raster maps for VP and VS2 can be combined by applying arithmetic operators. Before proceeding it is important to show raster maps (Figure 3) computed directly from new variables.

The first operation is the sum:  $VP+VS2$ . Results of this operation are presented in Figure 4. The result (Figure 4a) is simply the sum of raster maps (for VP and VS2). The variance of the sum (Figure 4b) is equal to the sum of variances for VP ( $\sigma_x^2$ ) and VS2 ( $\sigma_y^2$ ) plus twice the covariance between VP and VS2 ( $\sigma_{xy}$ ). The variance of the sum has great influence of the variable VP and the covariance as well. Results illustrated in Figure 4 are exactly the same as shown in Figure 3a.

Results of the subtraction of VS2 from VP are presented in Figure 5. The mathematical expectation and variance follow the same raster operations as for the sum. However, the variance is equal to the sum of variances minus twice the covariance between VP and VS2. The variance of the difference between random variables VP and VS2 shows a reasonable correlation with the variance of VP, but almost no correlation

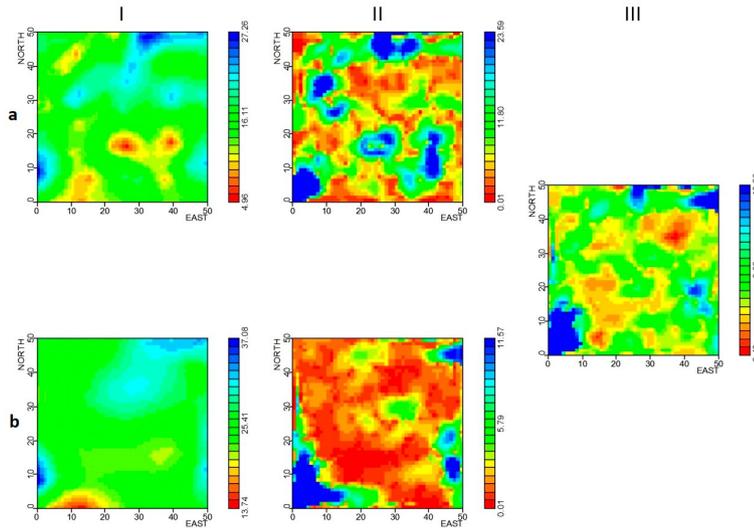
**TABLE III**  
**Mean and variance for combined random variables.**

$f(x,y)$	Mean	Variance
$X + Y$	$15.56775 + 25.05096 = 40,61871$ (10)	$21.12972 + 14.23615 + 26.12052 = 61.48639$ (11)
$X - Y$	$15.56775 - 25.05096 = -9.48321$ (10)	$21.12972 + 14.23615 - 26.12052 = 9.24535$ (11)
$XY$	$15.56775 * 25.05096 + 13.06026 = 403.04734$ (12)	$13259.97187 + 3450.19941 + 636.64517 - 170.57033 + 10186.66456 + 729.07689 + 479.85943 = 28571.847$ (13)
$X/Y$	$0.62144 - 0.02081 + 0.01410 = 0.61473$ (14)	$0.03367 + 0.00876 - 0.02587 = 0.01656$ (16)
	$0.62144 - 0.02081 + 0.01410 + 0.0009256476 - 0.0004044825 = 0.61525$ (15)	$0.03367 + 0.00876 - 0.02587 + 0.00118 + 0.00060 - 0.00185 + 0.00122 + 0.00122 - 0.00050 - 0.00151 = 0.01692$ (17)

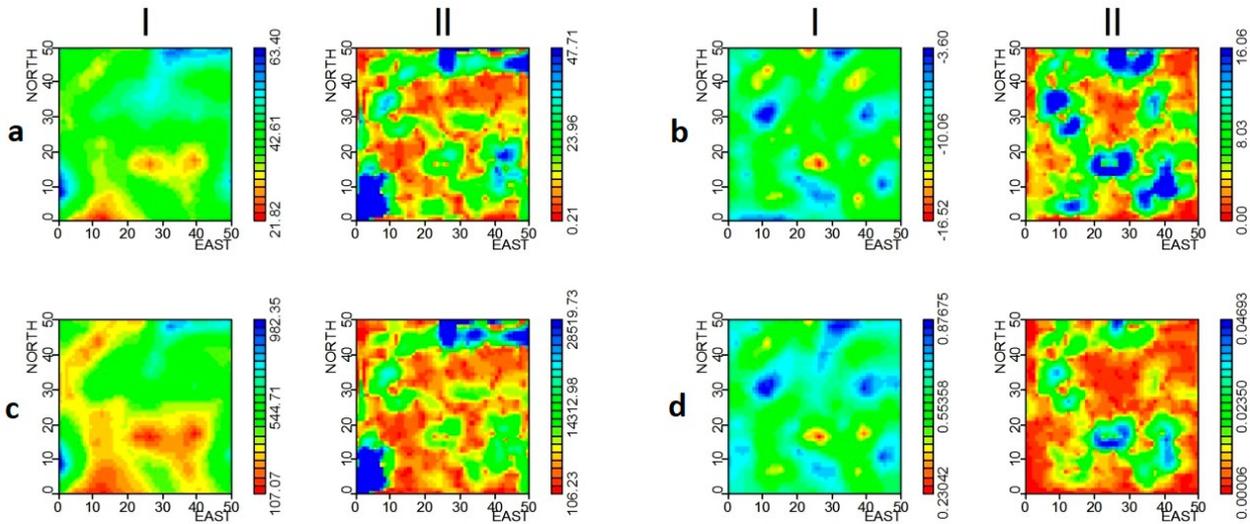
with the variance of VS2 and the covariance. Mean and variances computed after equations (10) and (11) for addition and subtraction match those values computed directly from equations (21) and (23). Once again, results can be compared graphically with images presented in Figure 3b.

The results of the multiplication of VP and VS2 are presented in Figure 6, which can be compared with results of direct multiquadric interpolation of the variable VPxVS2 (Figure 3c). The mathematical expectation of this operation  $E[f(x,y)]$  is equal to the product of VP and VS2 plus the covariance between VP and VS2. Note that the mathematical expectation has the contribution of the covariance.  $E[f(x,y)]$  is highly correlated with the means  $\mu_x$  and  $\mu_y$  and low correlation with the covariance. The uncertainty  $Var[f(x,y)]$  is equal to term 1 ( $\mu_y^2 \sigma_x^2$ ) plus term 2 ( $\mu_x^2 \sigma_y^2$ ) plus term 3 ( $E[(x - \mu_x)^2 (y - \mu_y)^2]$ ) minus term 4 ( $\sigma_{xy}^2$ ) plus term 5 ( $2\mu_x \mu_y \sigma_{xy}$ ) plus term 6 ( $2\mu_y E[(x - \mu_x)^2 (y - \mu_y)]$ ) plus term 7 ( $2\mu_x E[(x - \mu_x) (y - \mu_y)^2]$ ). Next, the variance was calculated after second-order Taylor expansion (equation (13)) and therefore higher order moments have been included. Equations (12) and (13) provided the same results as applying equations (21) and (23) for the new random variable VPxVS2.

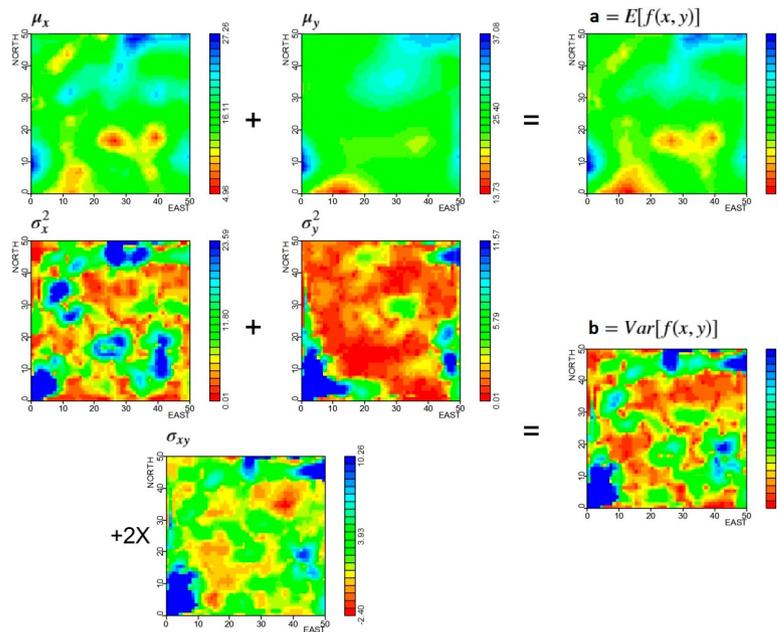
The calculation of mean and variance for the quotient of VP and VS2 is more complex since the function  $X/Y$  is infinitely differentiable with respect to  $Y$ . The mathematical expectation is not only equal to the quotient between VP and VS2 ( $\mu_x/\mu_y$ ), but includes two other terms:  $\sigma_{xy}/\mu_y^2$  and  $\mu_x \sigma_y^2/\mu_y^3$  (Figure 7a). This is the result after equation (14), which results from second-order Taylor expansion. Results from application of the equation (15) are shown in Figure 7b. Although results look very similar to each other, the mean according to equation (15) is more accurate than that provided by equation (14). Figures 7a, b basically display the same pattern as shown in Figure 3d-I. For calculating the variance of the ratio of X to Y, we have equation (16) that results from first-order Taylor expansion and another equation (17) derived after second-order Taylor expansion (Figure 8). It is virtually impossible to see the difference between results of these approaches in Figure 8. Thus, means and variances can be compared by being displayed in a scatter plot (Figure 9). Although, mean values are very close to each other, we can see that equation (15) provides



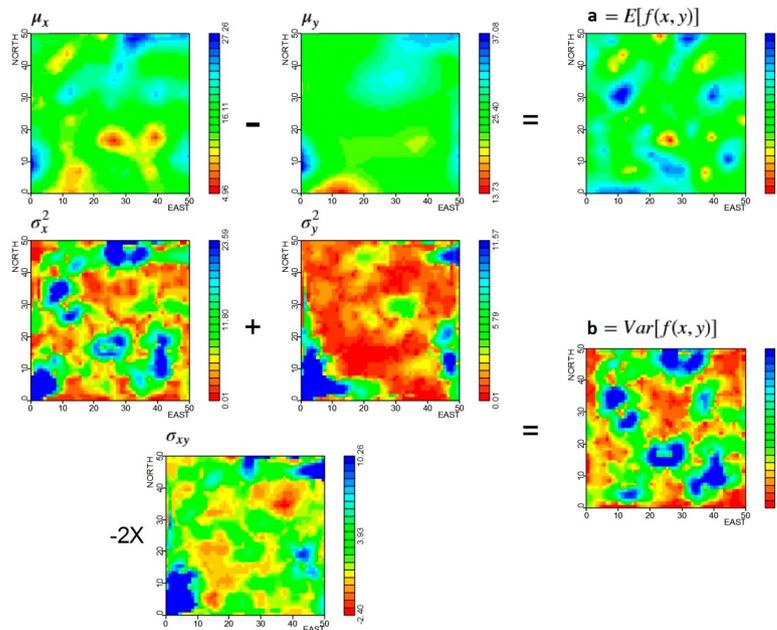
**Figure 2** - Raster maps for variables VP (a) and for variable VS2 (b). Interpolated variables in column I, uncertainties in column II and covariance between VP and VS2 in column III. Interpolation variances and covariance were displayed for a maximum equal to 95% of their distributions (for the interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).



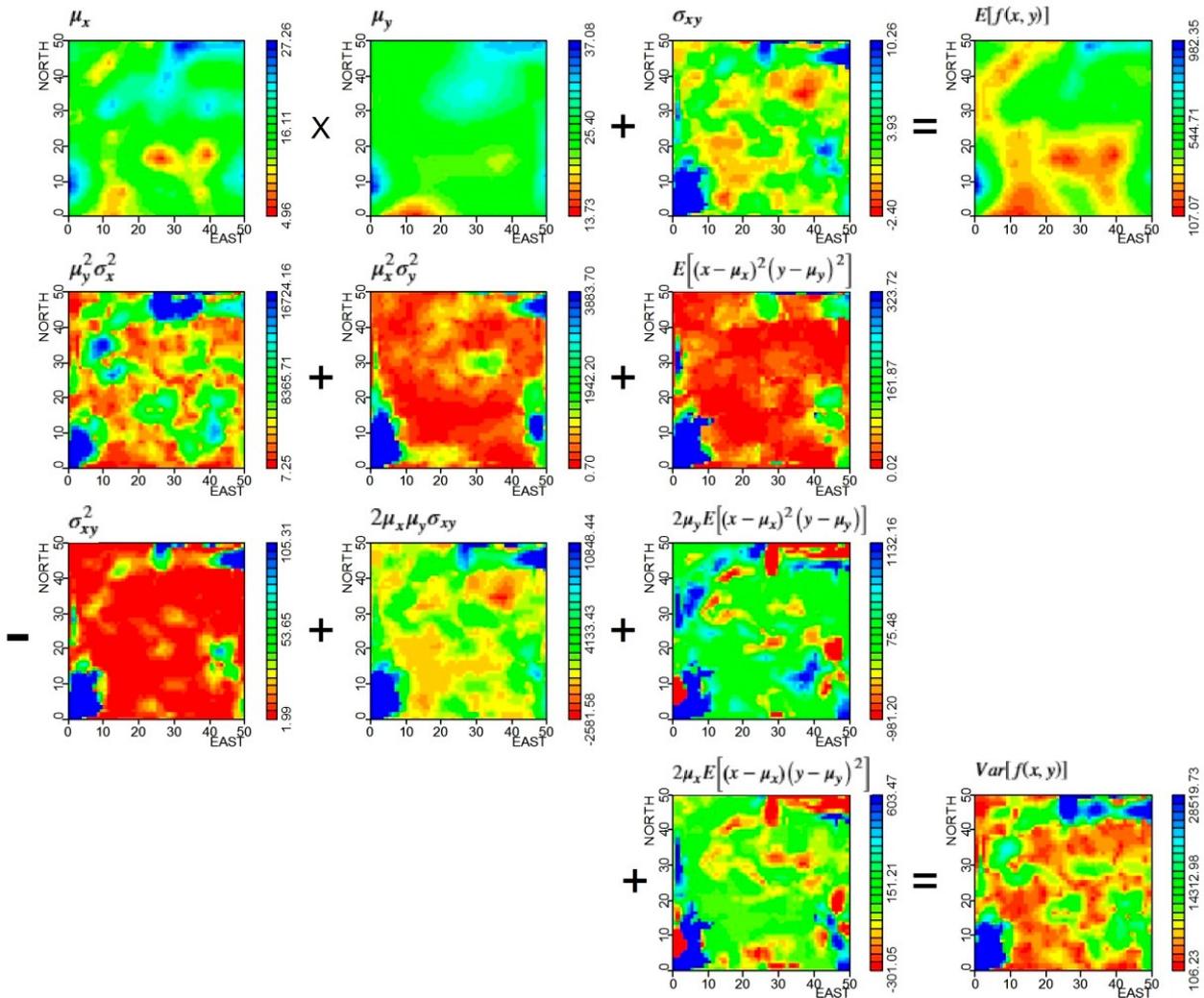
**Figure 3** - Raster maps computed directly from random variables VP+VS2 (a); VP-VS2 (b); VPxVS2 (c) and VP/VS2 (d). I) Interpolated variables; II) variances (for the interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).



**Figure 4** - Results of arithmetic operation VP+VS2. The result (a) is equal to the sum of VP ( $\mu_x$ ) and VS2 ( $\mu_y$ ). The associated uncertainty (b) is equal to the variance of VP ( $\sigma_x^2$ ) plus variance of VS2 ( $\sigma_y^2$ ) plus twice the covariance between VP and VS2 ( $\sigma_{xy}$ ) (for the interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).



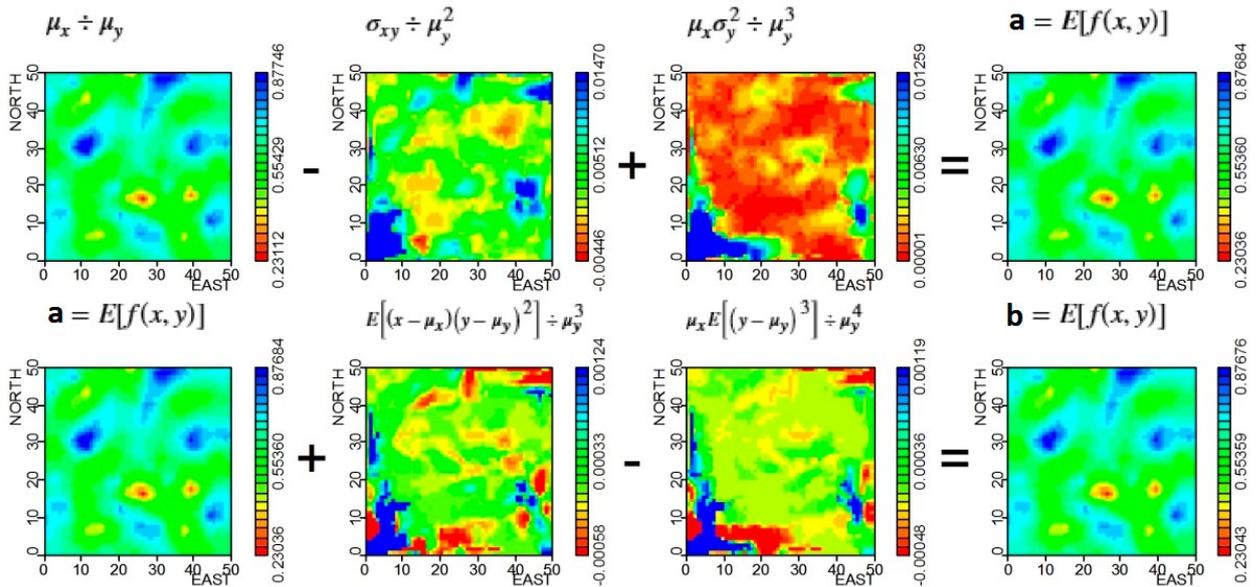
**Figure 5** - Results of arithmetic operation VP-VS2. The result (a) is equal to the sum of VP ( $\mu_x$ ) and VS2 ( $\mu_y$ ). The associated uncertainty (b) is equal to the variance of VP ( $\sigma_x^2$ ) plus variance of VS2 ( $\sigma_y^2$ ) minus twice the covariance between VP and VS2 ( $\sigma_{xy}$ ) (for the interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).



**Figure 6** - Results of arithmetic operation VPxVS2. The result ( $E[f(x,y)]$ ) is equal to the product of VP ( $\mu_x$ ) and VS2 ( $\mu_y$ ) plus the covariance between VP and VS2 ( $\sigma_{xy}$ ). The associated uncertainty ( $Var[f(x,y)]$ ) is equal to the term 1 ( $\mu_y^2 \sigma_x^2$ ) plus the term 2 ( $\mu_x^2 \sigma_y^2$ ) plus the term 3 ( $E[(x - \mu_x)^2 (y - \mu_y)^2]$ ) minus the term 4 ( $\sigma_{xy}^2$ ) plus the term 5 ( $2\mu_x \mu_y \sigma_{xy}$ ) plus the term 6 ( $2\mu_y E[(x - \mu_x)^2 (y - \mu_y)]$ ) plus the term 7 ( $2\mu_x E[(x - \mu_x)(y - \mu_y)^2]$ ) (for the interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).

a better approach than the equation (14). On the other hand, the variance after equation (17) is much better than that provided by equation (16).

We can follow arithmetic operations taking as an example the cell with coordinates CX=29.50 and CY=23.5 (Table IV). With eight neighbor data we interpolated this cell using field data for VP and VS2. All other variables (VP+VS2, VP-VS2, VPxVS2 and VP/VS2) were derived from these measurements (VP and VS2). Note that we can compute mean and variance directly for derived variables and thus these values can be checked against statistics computed using formulas listed in Table I. In Table IV, the data point with coordinates (CX=26.5, CY=25.5) has a weight equal to zero. This is because we applied an algorithm for correcting negative weights (Rao and Journel 1997), in which a constant equal to the modulus of the largest negative weight was added to all weights and then restandardized to a sum equal to one.

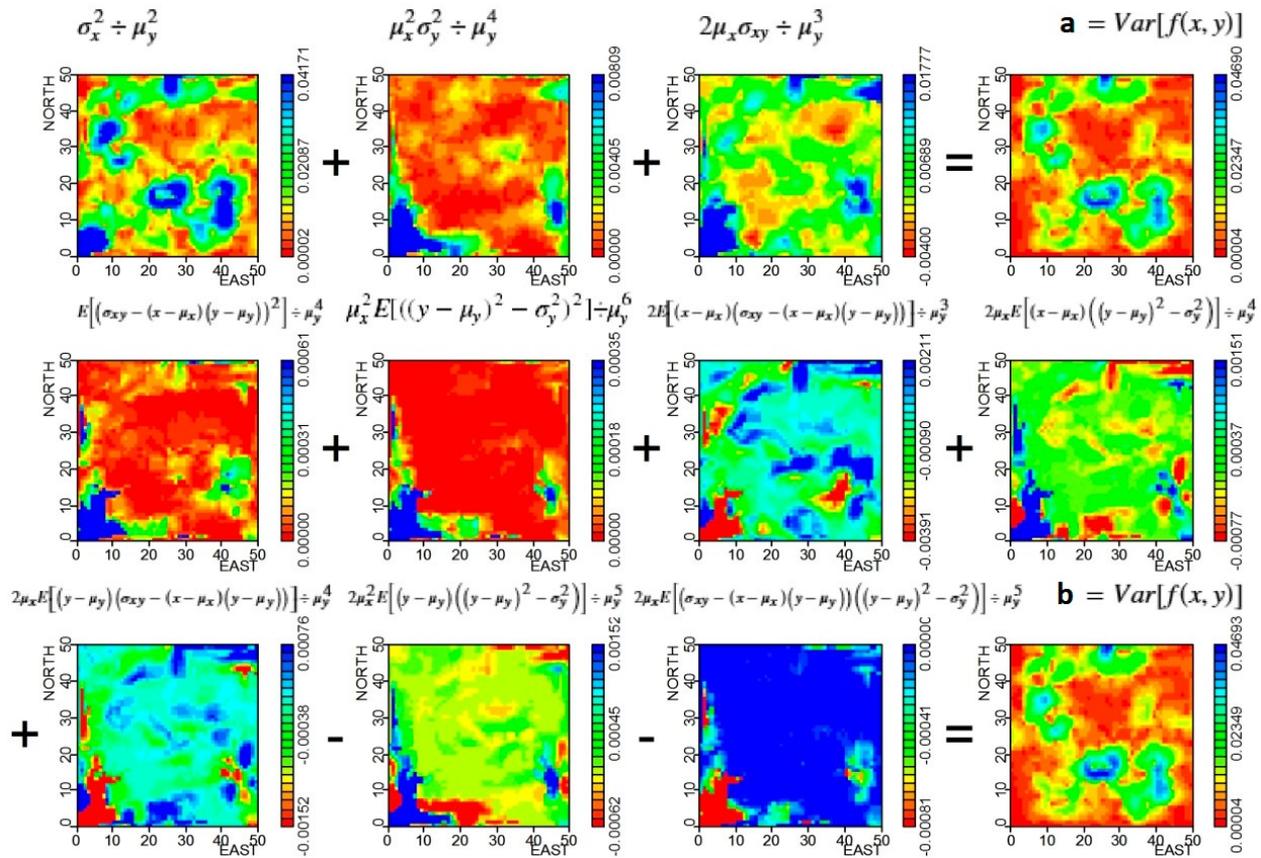


**Figure 7** - Mean values for arithmetic operation VP/VS2. (a - upper row) mean after equation (14); (b - lower row) mean according to equation (15) (for the interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).

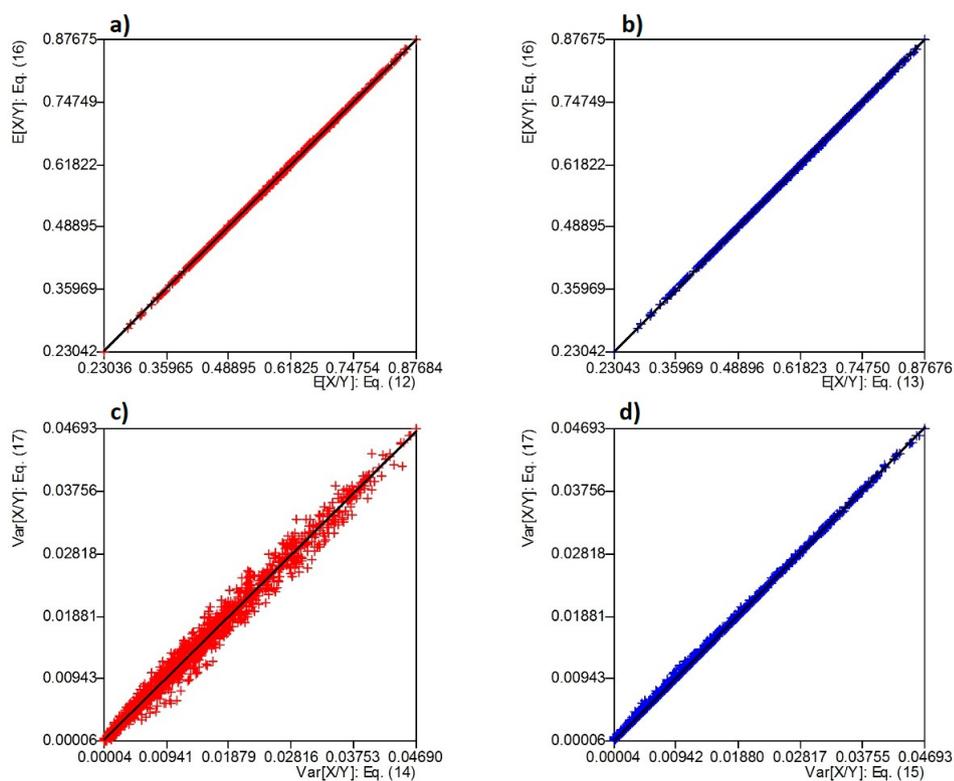
Mean and variances for the cell with coordinates (CX=29.50, CY=23.5) after combining random variables VP and VS2 are presented in Table V. As can be seen, statistics for VP+VS2, VP-VS2 and VPxVS2 match with statistics computed directly from derived variables listed in Table IV. However, for VP/VS2 we obtain only approximate results.

### CONCLUSIONS

Propagation of uncertainty coming from arithmetic operations between random variables is well known in statistics (Mood et al. 1974). Heuvelink et al. (1989) and Heuvelink (1998) have established the method of calculating mean and variances of the output raster map from several input maps based on Taylor method. This paper showed that propagation of uncertainty depends on a reliable measure of local accuracy and local covariance. We proved that it is possible to get exact mean and variance of the output raster map for map operations involving addition, subtraction and multiplication. For the mean of quotient, we proposed to



**Figure 8** - Variances of the arithmetic operation VP/VS2. **a)** resulting variance from first order Taylor expansion (equation (15)); **b)** variance calculated using second order Taylor expansion (equation (17)) (for the interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).



**Figure 9** - Scatterplots for means of the ratio of X to Y: **a** = equation (14) and **b** = equation (15); variances of the ratio: **c** = equation (16) and **d** = equation (17) (see the colors in the online version).

**TABLE IV**  
**Neighboring data around a cell with coordinates (CX=29.50, CY=23.5), mean and variances for variables VP, VS2, VP+VS2, VP-VS2, VPxVS2 and VP/VS2.**

CX	CY	W	VP	VS2	VP+VS2	VP-VS2	VPxVS2	VP/VS2
30.50	26.50	0.227	13.795	25.726	39.521	-11.931	354.890	0.53623
37.50	26.50	0.018	16.141	24.986	41.127	-8.845	403.299	0.64600
26.50	25.50	0	17.309	26.100	43.409	-8.791	451.765	0.66318
28.50	23.50	0.573	15.470	24.568	40.038	-9.098	380.067	0.62968
26.50	16.50	0.030	4.049	21.438	25.487	-17.389	86.802	0.18887
24.50	21.50	0.021	12.915	23.497	36.412	-10.581	303.464	0.54964
32.50	18.50	0.043	13.347	21.240	34.587	-7.894	283.490	0.62839
33.50	20.50	0.088	15.054	21.954	37.008	-6.900	33.496	0.68571
$Z_{MQ}^*(x_o)$			14.577	24.349	38.926	-9.773	355.831	0.59867
$S_o^2(x_o)$			4.129	1.656	7.594	3.977	2901.194	0.00730

**TABLE V**  
**Mean and variance for combined random variables.**

$f(x,y)$	Mean	Variance
$X + Y$	$14.57656 + 24.34917 = 38.92573$ (10)	$4.12927 + 1.65640 + 2 * 0.90406 = 7.59379$ (11)
$X - Y$	$14.57656 - 24.34917 = -9.77261$ (10)	$4.12927 + 1.65640 - 2 * 0.90406 = 3.97755$ (11)
$XY$	$14.57656 * 24.34927 + 0.900406 = 355.83120$ (12)	$2448.17033 + 351.94621 + 29.44879 - 0.81733 + 641.75335 - 473.39666 - 95.91113 = 2901.19356$ (13)
$X/Y$	$\frac{14.57656}{24.34917} - \frac{0.90406}{(24.34917)^2} + \frac{14.57656 * 1.65640}{(24.34917)^3} = 0.59879$ (14)	$\frac{4.12927}{(24.34917)^2} + \frac{(14.57656)^2 * 1.65640}{(24.34917)^4} - \frac{(2 * 14.57656) * 0.90406}{(24.34917)^3} = 0.0061402737$ (16)
	$0.0061402737 + 0.0000814529 + 0.0000073085 + 0.0013467544 - 0.0002728552 - 0.0002728552 + 0.0001317723 - 0.0000235991 = 0.0071382523$ (17)	

compute it from a third-order Taylor expansion that provides a slightly better result than the mean computed from second-order Taylor expansion. For the variance of the ratio of X to Y, we demonstrated that the formula based on second-order Taylor expansion provides much better results than the first order.

#### ACKNOWLEDGMENTS

The authors wish to thank Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) for supporting this research.

#### REFERENCES

- ATKINSON PM AND FOODY GM. 2002. Uncertainty in remote sensing and GIS: fundamentals. In: Foody GM and Atkinson PM (Eds), *Uncertainty in Remote Sensing and GIS*, p. 1-18.
- BURROUGH PA. 1986. *Principles of geographical information system*. Oxford, Oxford University Press, 194 p.
- CROSETTO M, TARANTOLA S AND SALTELI A. 2000. Sensitivity and uncertainty analysis in spatial modelling based on GIS. *Agr Ecosyst Environ* 81: 71-79.
- FRISHMAN F. 1971. On the arithmetic means and variances of products and ratios of random variables. *National Technical Information Service, AD-785623*, p. 331-345.
- HARDY RL. 1971. Multiquadric equations of topography and other irregular surfaces. *J Geophys Res* 76: 1905-1915.
- HEUVELINK GB. 1998. *Error propagation in environmental modelling with GIS*. London, Taylor & Francis, 127 p.
- HEUVELINK GB, BURROUGH PA AND STEIN A. 1989. Propagation of errors in spatial modelling with GIS. *Int J Geogr Inf Syst* 3: 303-322.
- JOURNAL AG AND ROSSI M. 1989. When do we need a trend model in kriging? *Math Geol* 21: 715-739.
- MASKEY S AND GUINOT V. 2003. Improved first-order second moment method for uncertainty estimation in flood forecasting. *Hydrolog Sci J* 48: 183-196.
- MCCALLUM WG. 1998. *Multivariable calculus*. New York, J Wiley & Sons, 270 p.
- MOOD AM, GRAYBILL FA AND BOES DC. 1974. *Introduction to the theory of statistics*. Tokyo, McGraw Hill, 3<sup>rd</sup> ed., 564 p.
- RAO SE AND JOURNAL AG. 1997. Deriving conditional distributions from ordinary kriging. In: Baafi EY and Schofield NA (Eds), *Geostatistics Wollongong 96*: 92-102.

- WANG G, GERTNER GZ, FANG S AND ANDERSON AB. 2005. A methodology for spatial uncertainty analysis of remote sensing and GIS products *Photogramm Eng Rem S* 71: 1423-1432.
- WEIR MD AND HASS J. 2014. *Thomas' calculus – multivariable*. Boston, Person Education Inc, 13<sup>th</sup> ed., p. 572-1044.
- WELLMANN FJ, HOROWITZ GF, SCHILL E AND REGENAUER-LIEB K. 2010. Towards incorporating uncertainty of structural data in 3D geological inversion. *Tectonophysics* 490: 141-152.
- YAMAMOTO JK. 2000. An alternative measure of the reliability of ordinary kriging estimates. *Math Geol* 32: 489-509.
- YAMAMOTO JK. 2002. Ore reserve estimation using radial basis functions. *Rev Inst Geol* 23: 25-38.
- YAMAMOTO JK, MAO XM, KOIKE K, CROSTA AP, LANDIM PMB, HU HZ, WANG CY AND YAO LQ. 2012. Mapping an uncertainty zone between interpolated types of a categorical variable. *Comput Geosci* 40: 146-152.

## APPENDIX

### A1 - MEAN AND VARIANCE OF THE FUNCTION $f(x,y)$ USING TAYLOR EXPANSION

In this appendix, we develop equations for mean and variance of the function  $f(x,y)$  as presented in Table I. We will use Taylor expansion of  $f(x,y)$  to find the mean and variance of this function around mean values  $\mu_x$  and  $\mu_y$ . Besides mean values, variances of random variables ( $\sigma_x^2$  and  $\sigma_y^2$ ) and the covariance between random variables X and Y ( $\sigma_{xy}$ ) must be known. The development for the mean and variance of product and quotient of random variables was based on Taylor expansion, according to Mood et al. (1974).

According to Weir and Hass (2014), the general equation of Taylor's Formula for two variables at the point  $\theta = (a, b)$  is:

$$f(x,y) = f(\theta) + \left( h \frac{\partial f(\theta)}{\partial x} + k \frac{\partial f(\theta)}{\partial y} \right) + \sum_{i=2}^n \frac{1}{i!} \left( h \frac{\partial}{\partial x} + k \frac{\partial f}{\partial y} \right)^i f \Big|_{(\theta)} + \frac{1}{(n+1)!} \left( h \frac{\partial}{\partial x} + k \frac{\partial f}{\partial y} \right)^{n+1} f \Big|_{(a+ch, b+ck)} \quad (\text{A1.1})$$

Where  $h = (x - a)$ ,  $k = (y - b)$ ,  $0 < c < 1$  and the last term is the remainder.

Neglecting the remainder term and taking  $n=2$ , we have:

$$f(x,y) = f(a,b) + \frac{\partial f(a,b)}{\partial x} (x-a) + \frac{\partial f(a,b)}{\partial y} (y-b) + \frac{1}{2!} \left( \frac{\partial^2 f(a,b)}{\partial x^2} (x-a)^2 + 2 \frac{\partial^2 f(a,b)}{\partial x \partial y} (x-a)(y-b) + \frac{\partial^2 f(a,b)}{\partial y^2} (y-b)^2 \right) \quad (\text{A1.2})$$

Letting  $(a, b) = (\mu_x, \mu_y)$  in equation A1.2 we have:

$$f(x,y) = f(\mu_x, \mu_y) + f_x(\mu_x, \mu_y) (x - \mu_x) + f_y(\mu_x, \mu_y) (y - \mu_y) + \frac{1}{2!} \left( f_{xx}(\mu_x, \mu_y) (x - \mu_x)^2 + 2f_{xy}(\mu_x, \mu_y) (x - \mu_x) (y - \mu_y) + f_{yy}(\mu_x, \mu_y) (y - \mu_y)^2 \right) \quad (\text{A1.3})$$

Applying the expectation operator to expression A1.3:

$$E[f(x,y)] = E[f(\mu_x, \mu_y)] + E[f_x(\mu_x, \mu_y) (x - \mu_x)] + E[f_y(\mu_x, \mu_y) (y - \mu_y)] + \frac{1}{2!} \left( E[f_{xx}(\mu_x, \mu_y) (x - \mu_x)^2] + 2E[f_{xy}(\mu_x, \mu_y) (x - \mu_x) (y - \mu_y)] + E[f_{yy}(\mu_x, \mu_y) (y - \mu_y)^2] \right)$$

Finally:

$$E[f(x, y)] = f(\mu_x, \mu_y) + \frac{1}{2} \left( \sigma_x^2 f_{xx}(\mu_x, \mu_y) + 2\sigma_{xy} f_{xy}(\mu_x, \mu_y) + \sigma_y^2 f_{yy}(\mu_x, \mu_y) \right) \quad (\text{A1.4})$$

This is the general expression for mathematical expectation of  $f(x, y)$ . The variance of the function  $f(x, y)$  can be written as:

$$\text{Var}(f(x, y)) = E. \quad (\text{A1.5})$$

Replacing A1.2 and A1.3 in A1.4, we get:

$$\text{Var}(f(x, y)) = E. \quad (\text{A1.6})$$

*Calculation of mean and variance for sum and subtraction of random variables X and Y*

For  $f(x, y) = x \pm y$ , we have the following derivatives:

$$f_x(x, y) = 1$$

$$f_y(x, y) = \pm 1$$

$$f_{xx}(x, y) = 0$$

$$f_{xy}(x, y) = 0$$

$$f_{yy}(x, y) = 0$$

Applying them on equations A1.4 and A1.6, we have equations (10) and (11) of Table I:

$$E[f(x, y)] = f(\mu_x, \mu_y)$$

$$E[f(x, y)] = \mu_x \pm \mu_y \quad (\text{A1.7})$$

$$\text{Var}(f(x, y)) = E \left[ ((x - \mu_x) \pm (y - \mu_y))^2 \right]$$

$$\text{Var}(f(x, y)) = E \left[ (x - \mu_x)^2 \pm 2(x - \mu_x)(y - \mu_y) + (y - \mu_y)^2 \right]$$

$$\text{Var}(f(x, y)) = \sigma_x^2 + \sigma_y^2 \pm 2\sigma_{xy} \quad (\text{A1.8})$$

*Calculation of the mean and variance for the product of random variables X and Y*

For  $f(x, y) = xy$ , we have the following derivatives:

$$f_x(x, y) = y$$

$$f_y(x, y) = x$$

$$f_{xx}(x, y) = 0$$

$$f_{yy}(x, y) = 0$$

$$f_{xy}(x, y) = 1$$

Applying them on equation A1.4, we get equation ((12) - Table I):

$$E[f(x, y)] = \mu_x \mu_y + \sigma_{xy} \tag{A1.9}$$

Replacing derivatives in equation A1.6, we have:

$$Var(f(x, y)) = E \left\{ \left[ \mu_y(x - \mu_x) + \mu_x(y - \mu_y) + (x - \mu_x)(y - \mu_y) - \sigma_{xy} \right]^2 \right\} \tag{A1.10}$$

Notice that:

$$(a + b + c + d)^2 = a^2 + b^2 + c^2 + d^2 + 2ab + 2ac + 2ad + 2bc + 2bd + 2cd \tag{A1.11}$$

Expanding A1.10 according to A1.11 we obtain:

$$Var(f(x, y)) = E.$$

Applying definitions of variance, covariance and the general equation of variance A1.6, we get equation ((13) - Table I):

$$\begin{aligned} Var(f(x, y)) = & \mu_y^2 \sigma_x^2 + \mu_x^2 \sigma_y^2 + 2\mu_x \mu_y \sigma_{xy} + E \left[ (x - \mu_x)^2 (y - \mu_y)^2 \right] - (\sigma_{xy})^2 \\ & + 2\mu_y E \left[ (x - \mu_x)^2 (y - \mu_y) \right] + 2\mu_x E \left[ (x - \mu_x) (y - \mu_y)^2 \right] \end{aligned} \tag{A1.12}$$

*Calculation of the mean and variance for the ratio of X to Y*

For calculation of the mean of the quotient of random variables X and Y we will consider second and third order Taylor expansion as follows.

For function  $f(x, y) = x/y$ , we have the following derivatives:

$$\begin{aligned} f_x &= \frac{1}{y} & f_{xx} &= 0 & f_{xxx} &= 0 \\ & \rightarrow f_{xy} = \frac{-1}{y^2} \rightarrow & f_{xyy} &= \frac{2}{y^3} \\ f_y &= \frac{-x}{y^2} & f_{yyy} &= \frac{-6x}{y^4} \end{aligned}$$

Second order Taylor expansion for calculation of the mean of function  $f(x, y) = x/y$ :

Applying the derivatives (until second order) on the equation A1.4 we get (equation (14) - Table I):

$$E[f(x, y)] = \frac{\mu_x}{\mu_y} - \sigma_{xy}^2 \frac{1}{\mu_y^2} + \sigma_y^2 \frac{\mu_x}{\mu_y^3} \quad (\text{A1.13})$$

Third order Taylor expansion for calculation of the mean of function  $f(x, y) = x/y$ :

Developing equation A1.1 for  $n=3$ , we get:

$$\begin{aligned} f(x, y) = & f(a, b) + \frac{\partial f(a, b)}{\partial x} (x - a) + \frac{\partial f(a, b)}{\partial y} (y - b) \\ & + \frac{1}{2!} \left( \frac{\partial^2 f(a, b)}{\partial x^2} (x - a)^2 + 2 \frac{\partial^2 f(a, b)}{\partial x \partial y} (x - a)(y - b) + \frac{\partial^2 f(a, b)}{\partial y^2} (y - b)^2 \right) \\ & + \frac{1}{3!} \left( \frac{\partial^3 f(a, b)}{\partial x^3} (x - a)^3 + 3 \frac{\partial^3 f(a, b)}{\partial x^2 \partial y} (x - a)^2 (y - b) + 3 \frac{\partial^3 f(a, b)}{\partial x \partial y^2} (x - a)(y - b)^2 + \frac{\partial^3 f(a, b)}{\partial y^3} (y - b)^3 \right) \end{aligned} \quad (\text{A1.14})$$

Letting  $(a, b) = (\mu_x, \mu_y)$  in equation A1.14, we have:

$$\begin{aligned} f(x, y) = & f(\mu_x, \mu_y) + f_x(\mu_x, \mu_y)(x - \mu_x) + f_y(\mu_x, \mu_y)(y - \mu_y) \\ & + \frac{1}{2!} (f_{xx}(\mu_x, \mu_y)(x - \mu_x)^2 + 2f_{xy}(\mu_x, \mu_y)(x - \mu_x)(y - \mu_y) + f_{yy}(\mu_x, \mu_y)(y - \mu_y)^2) \\ & + \frac{1}{3!} (f_{xxx}(\mu_x, \mu_y)(x - \mu_x)^3 + 3f_{xxy}(\mu_x, \mu_y)(x - \mu_x)^2(y - \mu_y) + 3f_{xyy}(\mu_x, \mu_y)(x - \mu_x)(y - \mu_y)^2 \\ & + f_{yyy}(\mu_x, \mu_y)(y - \mu_y)^3) \end{aligned} \quad (\text{A1.15})$$

Replacing derivatives in equation A1.15 we have (equation (15) - Table I):

$$E[f(x, y)] = \frac{\mu_x}{\mu_y} - \sigma_{xy}^2 \frac{1}{\mu_y^2} + \sigma_y^2 \frac{\mu_x}{\mu_y^3} + \frac{E[(x - \mu_x)(y - \mu_y)^2]}{\mu_y^3} - \frac{\mu_x}{\mu_y^4} E[(y - \mu_y)^3] \quad (\text{A1.16})$$

First order Taylor expansion for calculation of the variance of function  $f(x, y) = x/y$ :

To obtain the expression for variance using First order Taylor expansion, we will develop the equation A1.1 without the summatory term, then we have:

$$f(x, y) = f(a, b) + \frac{\partial f(a, b)}{\partial x} (x - a) + \frac{\partial f(a, b)}{\partial y} (y - b) \quad (\text{A1.17})$$

Letting  $(a, b) = (\mu_x, \mu_y)$ , we have:

$$f(x, y) = f(\mu_x, \mu_y) + f_x(\mu_x, \mu_y)(x - \mu_x) + f_y(\mu_x, \mu_y)(y - \mu_y)$$

Then, we will get for the expectation:

$$E[f(x, y)] = f(\mu_x, \mu_y)$$

And for the variance:

$$\begin{aligned} \text{Var}(f(x, y)) &= E \left[ (f_x(\mu_x, \mu_y)(x - \mu_x) + f_y(\mu_x, \mu_y)(y - \mu_y))^2 \right] \\ \text{Var}(f(x, y)) &= E \left[ f_x^2(\mu_x, \mu_y)(x - \mu_x)^2 + 2f_x(\mu_x, \mu_y)f_y(\mu_x, \mu_y)(x - \mu_x)(y - \mu_y) + f_y^2(\mu_x, \mu_y)(y - \mu_y)^2 \right] \end{aligned}$$

Applying the corresponding derivatives (until first order) we get (equation (16) - Table I):

$$\text{Var}(f(x, y)) = \frac{\sigma_x^2}{\mu_y^2} - 2\frac{\mu_x}{\mu_y^3}\sigma_{xy} + \frac{\mu_x^2}{\mu_y^4}\sigma_y^2 \tag{A1.18}$$

Second order Taylor expansion for calculation of the variance of function  $f(x, y) = x/y$  (equation A1.6):

$$\begin{aligned} \text{Var}(f(x, y)) &= E \left\{ \left[ \frac{1}{\mu_y}(x - \mu_x) - \frac{\mu_x}{\mu_y^2}(y - \mu_y) - \frac{-1}{\mu_y^2}[(x - \mu_x)(y - \mu_y) - \sigma_{xy}] + \frac{\mu_x}{\mu_y^3}[(y - \mu_y)^2 - \sigma_y^2] \right]^2 \right\} \\ \text{Var}(f(x, y)) &= E \left\{ \left[ \frac{(x - \mu_x)}{\mu_y} - \frac{\mu_x}{\mu_y^2}(y - \mu_y) + \frac{[\sigma_{xy} - (x - \mu_x)(y - \mu_y)]}{\mu_y^2} + \frac{\mu_x}{\mu_y^3}[(y - \mu_y)^2 - \sigma_y^2] \right]^2 \right\} \end{aligned} \tag{A1.19}$$

Expanding A1.19 according to A1.11 we obtain:

$$\text{Var}(f(x, y)) = E. \tag{A1.20}$$

Developing and simplifying, we get equation ((17) - Table I):

$$\begin{aligned} \text{Var}[f(x, y)] &\approx \frac{\sigma_x^2}{\mu_y^2} + \frac{\mu_x^2}{\mu_y^4}\sigma_y^2 - 2\frac{\mu_x}{\mu_y^3}\sigma_{xy} + \frac{1}{\mu_y^4}E[(\sigma_{xy} - (x - \mu_x)(y - \mu_y))^2] + \frac{\mu_x^2}{\mu_y^6}E\left[\left((y - \mu_y)^2 - \sigma_y^2\right)^2\right] \\ &+ \frac{2}{\mu_y^3}E[(x - \mu_x)[\sigma_{xy} - (x - \mu_x)(y - \mu_y)]] + 2\frac{\mu_x}{\mu_y^4}E[(x - \mu_x)[(y - \mu_y)^2 - \sigma_y^2]] \\ &- 2\frac{\mu_x}{\mu_y^4}E[(y - \mu_y)[\sigma_{xy} - (x - \mu_x)(y - \mu_y)]] - 2\frac{\mu_x^2}{\mu_y^5}E[(y - \mu_y)[(y - \mu_y)^2 - \sigma_y^2]] \\ &+ 2\frac{\mu_x}{\mu_y^5}E\left[[\sigma_{xy} - (x - \mu_x)(y - \mu_y)][(y - \mu_y)^2 - \sigma_y^2]\right] \end{aligned} \tag{A1.21}$$

A2 - HEUVELINK'S APPROACH FOR CALCULATION OF THE MEAN AND VARIANCE OF FUNCTION  $f(x, y)$  BASED ON SECOND ORDER TAYLOR EXPANSION

In this appendix we wanted to show that the traditional approach used in Appendix A1 gives the same results as obtained by Heuvelink et al. (1989) and Heuvelink (1998). Actually, these references presented a general formula of second order Taylor expansion for n random variables.

The second order Taylor expansion of a function with n variables at the point  $M=(\mu_1, \dots, \mu_n)$  and neglecting the remainder can be written as (Heuvelink et al. 1989):

$$f(\hat{x}) = f(M) + \sum_{i=1}^n [(x_i - \mu_i) f_{x_i}(M)] + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n [(x_i - \mu_i)(x_j - \mu_j) f_{x_i x_j}(M)] \quad (\text{A2.1})$$

Then, the mean around M is computed as (Heuvelink 1989):

$$\begin{aligned} E(f(M)) &= E[f(M)] + E \left[ \sum_{i=1}^n [(x_i - \mu_i) f_i(M)] \right] + \frac{1}{2} E \left[ \sum_{i=1}^n \sum_{j=1}^n [(x_i - \mu_i)(x_j - \mu_j) f_{ij}(M)] \right] \\ E(f(M)) &= f(M) + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n [\rho_{ij} \sigma_i \sigma_j f_{ij}(M)] \end{aligned} \quad (\text{A2.2})$$

Where  $\rho_{ij}$  is the correlation coefficient between variables  $i$  and  $j$ ;  $\sigma_i$  is the standard deviation for variable  $i$  and  $\sigma_j$  is the standard deviation for variable  $j$ .

$$\begin{aligned} \text{Var}(\hat{x}) &= E \left\{ \left[ \sum_{k=1}^n [(x_k - \mu_k) f_k(M)] + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n [(x_i - \mu_i)(x_j - \mu_j) - \rho_{ij} \sigma_i \sigma_j] f_{ij}(M) \right]^2 \right\} \\ \text{Var}(\hat{x}) &= \sum_{k=1}^n \sum_{l=1}^n \rho_{kl} \sigma_k \sigma_l f_k(M) f_l(M) \\ &+ \sum_{k=1}^n \sum_{i=1}^n \sum_{j=1}^n \left\{ E [(x_k - \mu_k) ((x_i - \mu_i)(x_j - \mu_j) - \rho_{kl} \sigma_k \sigma_l)] f_k(M) f_{ij}(M) \right\} \\ &+ \frac{1}{4} \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n \sum_{l=1}^n \left\{ E [(x_i - \mu_i)(x_j - \mu_j)(x_k - \mu_k)(x_l - \mu_l) - \rho_{ij} \sigma_i \sigma_j \rho_{kl} \sigma_k \sigma_l] f_{ij}(M) f_{kl}(M) \right\} \end{aligned} \quad (\text{A2.3})$$

Notice that for  $n=2$ , and setting  $x_1=x$  and  $x_2=y$ , the expression A2.2 becomes:

$$E(f(x,y)) = f(\mu_x, \mu_y) + \frac{1}{2} \left( \sigma_x^2 f_{xx}(\mu_x, \mu_y) + 2\sigma_{xy} f_{xy}(\mu_x, \mu_y) + \sigma_y^2 f_{yy}(\mu_x, \mu_y) \right) \quad (\text{A2.4})$$

Expressions A1.3 and A2.4 are the same, so the development of the expectation for an arithmetic operation is the same as shown in Appendix A1. Therefore, equations (10), (12) and (14) of Table I can be derived from A2.4.

For the variance, we will need to expand the expression A2.3 and consider the arithmetic operation regarding the function  $f(x,y)$ . Then at the point  $\mu = (\mu_x, \mu_y)$  we have:

$$\begin{aligned} \text{Var}(x,y) &= \sigma_x^2 f_x^2(\mu) + 2\sigma_{xy} f_x(\mu) f_y(\mu) + \sigma_y^2 f_y^2(\mu) \\ &+ E \left[ (x - \mu_x)^3 - (x - \mu_x) \sigma_x^2 \right] f_x(\mu) f_{xx}(\mu) \\ &+ 2E \left[ (x - \mu_x)^2 (y - \mu_y) - (x - \mu_x) \sigma_{xy} \right] f_x(\mu) f_{xy}(\mu) \\ &+ E \left[ (x - \mu_x) (y - \mu_y)^2 - (x - \mu_x) \sigma_y^2 \right] f_x(\mu) f_{yy}(\mu) \\ &+ E \left[ (x - \mu_x)^2 (y - \mu_y) - (y - \mu_y) \sigma_x^2 \right] f_y(\mu) f_{xx}(\mu) \\ &+ 2E \left[ (x - \mu_x) (y - \mu_y)^2 - (y - \mu_y) \sigma_{xy} \right] f_y(\mu) f_{xy}(\mu) \\ &+ E \left[ (y - \mu_y)^3 - (y - \mu_y) \sigma_y^2 \right] f_y(\mu) f_{yy}(\mu) + \frac{1}{4}. \end{aligned}$$

*Calculation of the variance for sum and subtraction of random variables X and Y*

Since for  $f(x, y) = x \pm y$  all second order derivatives are equal to zero, we have (equation (11) - Table I):

$$\text{Var}(x, y) = \sigma_x^2 + \sigma_y^2 \pm 2\sigma_{xy} \quad (\text{A2.5})$$

Because all second order derivatives are equal to zero.

*Calculation of the variance for the product of random variables X and Y*

$$\begin{aligned} \text{Var}(x, y) = & \sigma_x^2 \mu_y^2 + \sigma_y^2 \mu_x^2 + 2\sigma_{xy} \mu_x \mu_y + 2\mu_y E \left[ (x - \mu_x)^2 (y - \mu_y) \right] + 2\mu_x E \left[ (x - \mu_x) (y - \mu_y)^2 \right] \\ & + E \left[ (x - \mu_x)^2 (y - \mu_y)^2 \right] - (\sigma_{xy})^2 \end{aligned} \quad (\text{A2.6})$$

This is the same as equation ((13) - Table I).

*Calculation of the variance for the ratio of X to Y*

$$\begin{aligned} \text{Var}(x, y) = & \frac{\sigma_x^2}{\mu_y^2} - 2 \frac{\mu_x}{\mu_y^3} \sigma_{xy} + \frac{\mu_x^2}{\mu_y^4} \sigma_y^2 - \frac{2}{\mu_y^3} E \left[ (x - \mu_x)^2 (y - \mu_y) - (x - \mu_x) \sigma_{xy} \right] + \frac{2\mu_x}{\mu_y^4} E \left[ (x - \mu_x) (y - \mu_y)^2 - (x - \mu_x) \sigma_y^2 \right] \\ & + 2 \frac{\mu_x}{\mu_y^4} E \left[ (x - \mu_x) (y - \mu_y)^2 - (y - \mu_y) \sigma_{xy} \right] - \frac{2\mu_x^2}{\mu_y^5} E \left[ (y - \mu_y)^3 - (y - \mu_y) \sigma_y^2 \right] \\ & + \frac{1}{4} \left\{ -4 \frac{2\mu_x}{\mu_y^5} E \left[ (x - \mu_x) (y - \mu_y)^3 - \sigma_y^2 \sigma_{xy} \right] + 4 \frac{1}{\mu_y^4} E \left[ (x - \mu_x)^2 (y - \mu_y)^2 - (\sigma_{xy})^2 \right] + \frac{4\mu_x^2}{\mu_y^6} E \left[ (y - \mu_y)^4 - (\sigma_y^2)^2 \right] \right\} \\ \\ \text{Var}(x, y) = & \frac{\sigma_x^2}{\mu_y^2} + \frac{\mu_x^2}{\mu_y^4} \sigma_y^2 - 2 \frac{\mu_x}{\mu_y^3} \sigma_{xy} + \frac{2}{\mu_y^3} E \left[ (x - \mu_x) [\sigma_{xy} - (x - \mu_x) (y - \mu_y)] \right] \\ & + 2 \frac{\mu_x}{\mu_y^4} E \left[ (x - \mu_x) [(y - \mu_y)^2 - \sigma_y^2] \right] - 2 \frac{\mu_x}{\mu_y^4} E \left[ (y - \mu_y) [\sigma_{xy} - (x - \mu_x) (y - \mu_y)] \right] \\ & - 2 \frac{\mu_x^2}{\mu_y^5} E \left[ (y - \mu_y) [(y - \mu_y)^2 - \sigma_y^2] \right] - \frac{2\mu_x}{\mu_y^5} E \left[ (x - \mu_x) (y - \mu_y)^3 - \sigma_y^2 \sigma_{xy} \right] \\ & + \frac{E \left[ (x - \mu_x)^2 (y - \mu_y)^2 \right] - (\sigma_{xy})^2}{\mu_y^4} + \frac{\mu_x^2}{\mu_y^6} E \left[ (y - \mu_y)^4 - (\sigma_y^2)^2 \right] \end{aligned} \quad (\text{A2.7})$$

As we can see, some terms of equation A2.7 do not match A1.21. Actually, we can rewrite these terms as follows.

$$E \left[ [\sigma_{xy} - (x - \mu_x) (y - \mu_y)]^2 \right] = (\sigma_{xy})^2 - 2\sigma_{xy} E \left[ (x - \mu_x) (y - \mu_y) \right] + E \left[ (x - \mu_x)^2 (y - \mu_y)^2 \right] = E \left[ (x - \mu_x)^2 (y - \mu_y)^2 \right] - (\sigma_{xy})^2$$

$$E \left[ [(y - \mu_y)^2 - \sigma_y^2]^2 \right] = E \left[ (y - \mu_y)^4 \right] - 2\sigma_y^2 E \left[ (y - \mu_y)^2 \right] + (\sigma_y^2)^2 = E \left[ (y - \mu_y)^4 \right] - (\sigma_y^2)^2 = E \left[ (y - \mu_y)^4 - (\sigma_y^2)^2 \right]$$

$$\begin{aligned}
E \left[ [\sigma_{xy} - (x - \mu_x)(y - \mu_y)] [(y - \mu_y)^2 - \sigma_y^2] \right] &= E \left[ [(x - \mu_x)(y - \mu_y) - \sigma_{xy}] [\sigma_y^2 - (y - \mu_y)^2] \right] \\
&= E \left[ (x - \mu_x)(y - \mu_y) \sigma_y^2 \right] - E \left[ (x - \mu_x)(y - \mu_y)^3 \right] - E \left[ \sigma_{xy} \sigma_y^2 \right] \\
&\quad + E \left[ \sigma_{xy} (y - \mu_y)^2 \right] = E \left[ \sigma_{xy} \sigma_y^2 \right] - E \left[ (x - \mu_x)(y - \mu_y)^3 \right] \\
&= E \left[ \sigma_{xy} \sigma_y^2 - (x - \mu_x)(y - \mu_y)^3 \right]
\end{aligned}$$

Replacing these terms into equation A2.7 we get exactly the same as A1.21, that is the equation ((17) - Table I).