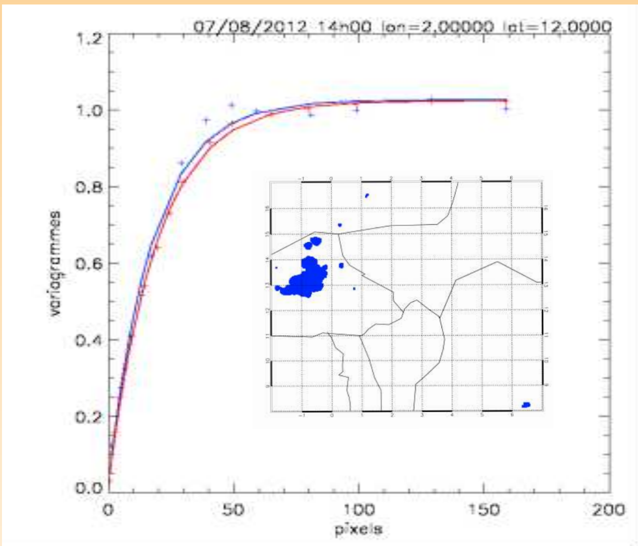
MTTM

Megha-Tropiques Technical Memorandum

Fast homogeneous and isotropic variogram computation using FFT for the TAPEER-BRAIN product

Nicolas Taburet, Philippe Chambon, Rémy Roca



megha-tropiques

January 2016

07

Fast homogeneous and isotropic variogram computation using FFT for the TAPEER-BRAIN product

Nicolas TABURET (LEGOS), Philippe CHAMBON (CNRM)
Rémy ROCA (LEGOS)

January 2016

1. INTRODUCTION

The TAPEER algorithm provides one degree/one day accumulated rainfall estimates and their associated sampling errors (Roca et al. 2010). The calculation of these errors, requires the estimate of characterisitic space and time decorellation lengths. Those quantities are computed as the e-folding distances of an exponential model that fits the spacial and temporal 10 day-averaged variograms (Lebel 1987, Chambon et al 2012). TAPEER rain/no rain indicator fields are used in order to avoid the effect of nearby field values noise in the determination of the variogram sills (Carr et al 1985).

This document details the approaches used to estimates the elementary variograms before performing any averaging over 10 days. The direct space method for spatial variogram estimation is presented as well as its potential limitations. A faster method based on Fast Fourier Transform (FFT) is then described and discussed.

2. DIRECT SPACE CALCULATION

2.1. Variograms

The variogram is defined as the variance of the difference between 2 field values :

$$2\gamma(\vec{h}) = \mathrm{Var}[V(\vec{x}) - V(\vec{x}+\vec{h})]$$

The empirical semi-variograms are therefore computed as :

$$\gamma(\vec{h}) = \frac{1}{2N(\vec{h})}\sum_{x\in A}(V(\vec{x}) - V(\vec{x}+\vec{h}))^2$$

As described in Chambon et al. (2012a), using an isotropy hypothesis the semi-variograms only depends on the inter-distance h. The square of the differences between all pairs of the indicator field (IF) are computed to obtain the empirical space variogram (figure 1 and equation (1)) :

$$(1)\ \gamma(h) = \frac{1}{2N(h)}\sum_{i=1}^{N(h)}[IF(\vec{x_1}) - IF(\vec{x_2})]^2$$

where N is the number of pairs separated by a lag h.

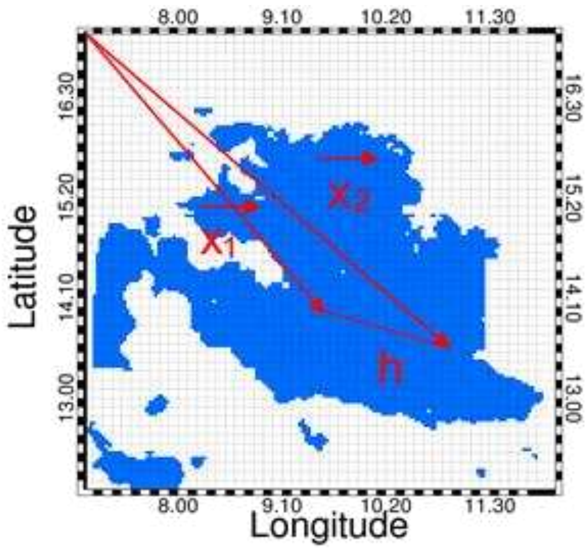


Fig.1 : Interdistance between a pair of two samples.

2.2. Limitations

It has been shown that edge effect can affect variogram calculations (Xu and Dowd, 2012). Using variogram calculation on simulated fields, they emphasized that the edge effect is negligible when the correlation distance a, defined as the lag at which the variogram plateau appears, is much smaller than the dimension L of the simulation. For a/L>0.3 noticeable effects are seen on the variograms, making their representation inacurate. Variograms are robust to the field size when a/L<0.2. With e-folding distances of the order of 50 km in West Africa, i.e. 95 % of the plateau value reached at 150 km, a 5°x5° window (approximately 500km) is close to the limit, the calculation will only be safe if the rain structure is at the center of the field and not close to its border.

The variograms computed on stationary fields (mean and variance do not vary significantly in space) satisfy the equation:

$$(2)\ \gamma(\vec{h}) = (\mathrm{cov}(V))(0) - (\mathrm{cov}(V))(\vec{h}) = \sigma^2 - (\mathrm{cov}(V))(\vec{h})$$

and therefore present a plateau. Most of the time, geostatistical fields are not stationary and therefore their variogram consist of the variogram stochastic component with the plateau and a deterministic trend (Ali, Lebel and Amani, 2002). Consequently, the variograms computed in direct space with equation (2) can present a trend at large scales.

3. VARIOGRAMS IN SPECTRAL SPACE

3.1. Rationale

For stationary fields, the semi-variogram function can be expressed as a function of the covariance of the field V as in equation (2). The covariance of the field can be written as follows, corresponding to a convolution product on the field V:

$$(3) \qquad C(h_x, h_y) = (\mathrm{cov}(V))(h_x, h_y) = \int_0^{L_x}\int_0^{L_y} V(x,y)V(x+h_x, y+h_y)\,\mathrm{dxdy}$$

A Fourier Transform on the covariance gives (4), a Fourier Transform of V and an inverse Fourier transform on the product in (4) should give consistent results as computing variograms in the physical space.

$$(4) \qquad \mathrm{FT}[C] = \mathrm{FT}[V]^{*}.\mathrm{FT}[V]$$

The bi-Fourier transform pair is:

$$(5) \qquad v_{\mathrm{mn}} = \frac{1}{L_x}\frac{1}{L_y}.\int_0^{L_x}\int_0^{L_y} V(\vec{h})\exp\left[-2\pi i\left(\frac{\mathrm{mx}}{L_x}+\frac{\mathrm{ny}}{L_y}\right)\right]\mathrm{dxdy}$$

$$(6) \qquad V(\vec{x}) = \int_{-M}^{+M}\int_{-N}^{+N} v_{\mathrm{mn}}\exp\left[2\pi i\left(\frac{\mathrm{mx}}{L_x}+\frac{\mathrm{ny}}{L_y}\right)\right]\mathrm{dmdn}$$

Equation (3) can be rewritten with the bi-Fourier coefficients:

$$(7) \qquad (\mathrm{cov}(V))(\vec{h}) = \int_{-M}^{+M}\int_{-N}^{+N}\int_{-M}^{+M}\int_{-N}^{+N} v_{\mathrm{mn}}v_{m'n'}\exp\left[2\pi i\left(\frac{\mathrm{mx}}{L_x}+\frac{\mathrm{ny}}{L_y}\right)\right]\exp\left[2\pi i\left(m'\frac{x+h_x}{L_x}+n'\frac{y+h_y}{L_y}\right)\right]\mathrm{dmdndm'dn'}$$

The homogeneity assumption (dependence on the separation vector $\vec{h}$ only) means that the spectral covariance $v_{\mathrm{mn}}v_{m'n'}$=$v_{\mathrm{mn}}v_{-m'-n'}$ must be null unless (m,n) and $(-m',-n')$ are equal (Berre, 2000). Hence:

$$(8) \qquad (\mathrm{cov}(V))(\vec{h}) = \int_{-M}^{+M}\int_{-N}^{+N} v_{\mathrm{mn}}v_{\mathrm{mn}}\exp\left[2\pi i\left(m\frac{h_x}{L_x}+n\frac{h_y}{L_y}\right)\right]\mathrm{dmdn}$$

A change of variable of the bi-Fourier transform can be performed into polar coordinates:

(9) $$(\mathrm{cov}(\boldsymbol{V}))(\vec{\boldsymbol{h}}) = \boldsymbol{L_x L_y} \int_{\boldsymbol{0}}^{\boldsymbol{K}} \int_{\boldsymbol{0}}^{\boldsymbol{2\pi}} \left|\boldsymbol{v}(\vec{\boldsymbol{k}})\right|^{\boldsymbol{2}} \exp\left[\boldsymbol{2\pi i}\frac{\boldsymbol{kh}}{\boldsymbol{D}}\cos(\vec{\boldsymbol{k}},\vec{\boldsymbol{h}})\right] \boldsymbol{k}\, \mathrm{d}k\boldsymbol{d\theta}$$

With $\vec{k} = (\frac{m}{L_x}, \frac{n}{L_y})$, $k = \sqrt{(\frac{m}{L_x})^2 + (\frac{n}{L_y})^2}$ and $\theta = \tan^{-1}(\frac{n}{L_y} / \frac{m}{L_x})$

The isotropy assumption (dependence on $\|\vec{h}\|$only) leads to $\left|v(\vec{k})\right|^2 = |v(k)|^2$and a separation of the integral over θ

(10) $$(\mathrm{cov}(\boldsymbol{V}))(\boldsymbol{h}) = \frac{\boldsymbol{L_x L_y}}{\boldsymbol{D^2}} \int_{\boldsymbol{0}}^{\boldsymbol{K}} |\boldsymbol{v}(\boldsymbol{k})|^{\boldsymbol{2}} \boldsymbol{k} \left(\int_{\boldsymbol{0}}^{\boldsymbol{2\pi}} \exp\left[\boldsymbol{2\pi i}\frac{\boldsymbol{kh}}{\boldsymbol{D}}\cos(\boldsymbol{k},\boldsymbol{h})\right] \boldsymbol{d\theta}\right) \mathrm{d}k$$

The integral over θ correspond to a zero-order Bessel function of the first kind, leading to Equation (11) for the expression of the homogeneous, isotropic covariance:

(11) $$(\mathrm{cov}(V))(h) = \frac{2\pi L_x L_y}{D^2} \int_0^K |v(k)|^2 J_O\left(+2\pi k \frac{h}{D}\right) k\, \mathrm{d}k$$

3.2. Implementation

The $|v(k)|$ coefficients are calculated using the isotropy assumption by averaging the Bi-Fourier coefficients obtained after Fast Fourier Transform of the rain/no rain indicator field. As presented in the following example, this isotropy assumption on the Bi-Fourier spectrum results in a smoothing of the latter (figure 3).

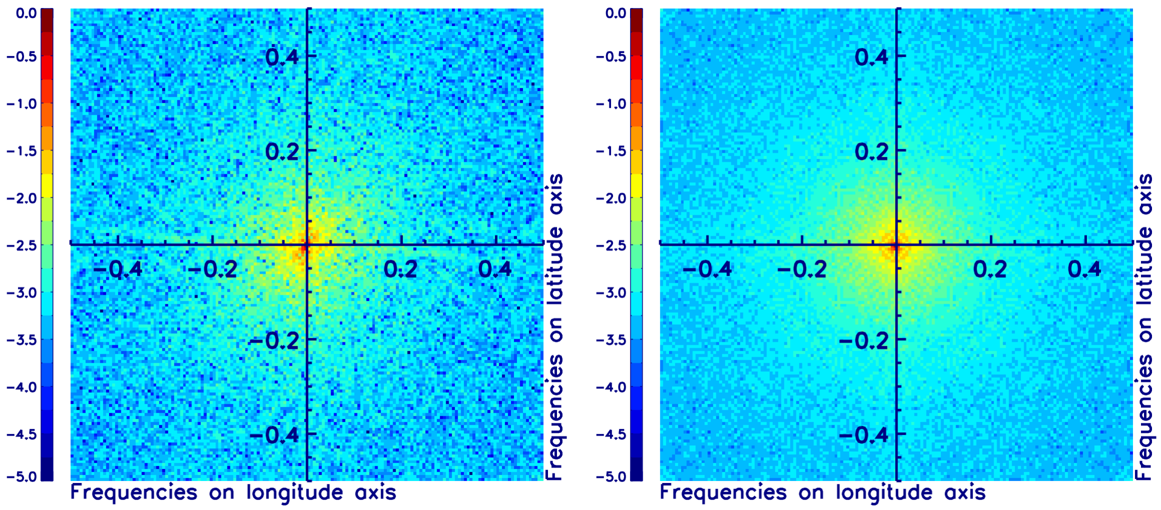


Fig.2 : left : base 10 logarithm of the Bi-Fourier spectrum resulting from the FFT computation on the indicator field. right : Bi-Fourier spectrum smoothed after isotropy assumption.

The integral over $k$in Equation (11) is computed using the trapezoidal rule. Depending on the value of $\|\vec{h}\|$, the zero-order Bessel function is characterized by rapid oscillations as function of k making the calculation sensitive to the discretization. Figure 4 represents three Bessel functions for $\|\vec{h}\| = 0.5$in black, $\|\vec{h}\| = 8.5$in red and $\|\vec{h}\| = 79.5$ in blue ($\|\vec{h}\|$values represents the separation between pixels). The frequencies at which the FFT coefficients are computed ensure enough sample of the Bessel function oscillations.

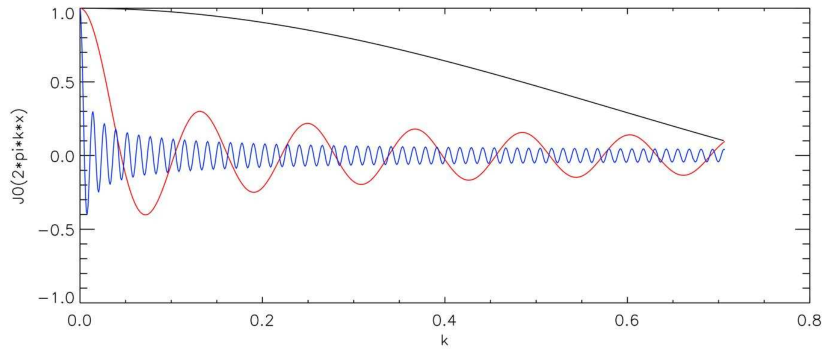


Fig.4 : Three Bessel J0 functions

The space variogram in each 5°x5° region is then obtain from the covariance and equation (2). The variograms obtained with the two methods are presented on figure 5 for the case of a field on which the direct method does not suffer from edge or trend effects.

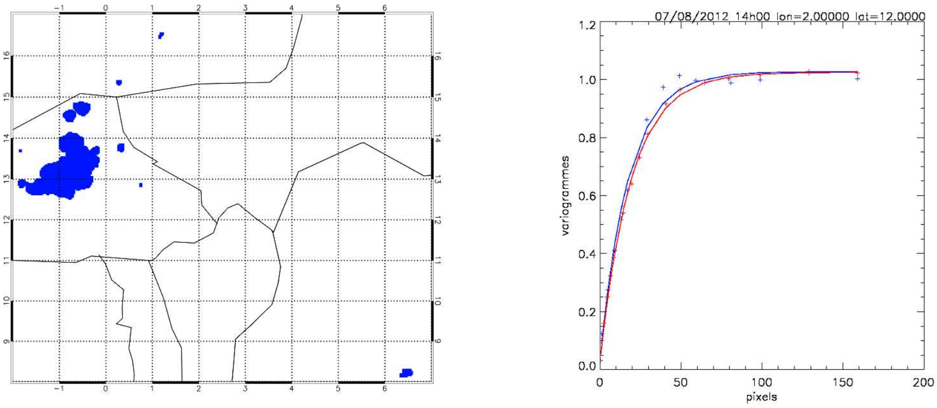


Fig.5 : Rain/ no rain field and the associated direct space (blue) and spectral space (red) variograms

3.3. Limitations

The FFT method does only presents the stochastic part of the variograms and therefore does not contain the 'trend' problem. The FFT method is also not sensitive to the position of the rainy structure within the field. Nevertheless the FFT method can still be affected by the edge effect if the 5°x5° field is not large enough compared to the characteristic size of the rainy structures. Detailed comparison of FFT and direct space estimations are discussed in further details in the ATBD document (Taburet et al. 2016).

4. CONCLUSIONS

The spectral method presents the advantage to provide variograms that do not depend on the situation of the rainy structure in the 5°x5° field. Those variograms also present the advantage of filtering the nugget as well as the long distance trend and providing empirical variograms with a plateau that are more adapted to be fitted with an exponential model that is used to obtain the characteristic size of rainy structures.

In term of CPU time, the spectral method divides that of the direct space method by a factor of 3. Since spatial variograms computation in the direct space over a decade requires approximately 10 hours of CPU time, the spectral method represents a significant improvement.

5. REFERENCES

Ali, Lebel and Amani, 2003, Invariance in the Spatial Structure of Sahelian Rain Fields at Climatological Scales, Journal of Hydrometeorology, vol 4 , 996-1010

Berre, Loïk, 2000: Estimation of Synoptic and Mesoscale Forecast Error Covariances in a Limited-Area Model. Mon. Wea. Rev., 128, 644–667. 10.1175/1520-0493(2000)128

Carr J., Bailey R., Deng E., 1985, Use of indicator variograms for an enhanced spatial analysis, Journal of the International Asociation for Mathematical Geology, 17 (8), 797-811.

Chambon P, Jobard I, Roca R, Viltard N., 2012(a). An investigation of the error budget of tropical rainfall accumulation derived from merged passive microwave and infrared satellite measurements. Q. J. R. Meteorol. Soc. 138: 000.000. DOI:10.1002/qj.1907

Chambon P., Roca R., Jobard I., Aublanc J., 2012(b), Megha-Tropiques Technical Memorandum n°4, TAPEER-BRAIN product: Algorithm Theorotical Basis Document

Chaoshui Xu and Peter A. Dowd, 2012, The Edge Effect in Geostatistical Simulations. Geostatistics Oslo 2012, pages 115-127

Lebel T., Bastin G., Obled C., Creutin J., 1987. On the accuracy of areal rainfall estimation, WATER RESOURCES RESEARCH, VOL. 23, NO. 11, PAGES 2123-2134

Marcotte, Denis, 1996: Fast variogram computation with FFT, Computers & Geosciences, Volume 22, Issue 10, December 1996, Pages 1175-1186, ISSN 0098-3004, 10.1016/S0098-3004(96)00026-X.

Roca, R., Chambon, P., Jobard, I., Kirstetter, P.-E., Gosset, ., and erg s, .-C. (2010). Comparing satellite and surface rainfall products over West Africa at meteorologically relevant scales during the AMMA campaign using error estimates. J. Appl. Meteorol. And Climatol., 49(4), 715–731.

Taburet N., Roca R., Chambon P., 2016, Megha Tropiques Note on the comparison of space variograms computed in direct or spectral space.

MTTM

Megha-Tropiques Technical Memorandum

Editorial committee :

Sophie Cloché

Michel Capderou

Laboratoire de Météorologie Dynamique (LMD / IPSL)

Ecole Polytechnique

F-91128 Palaiseau

France

Sophie.Bouffies-Cloche@ipsl.jussieu.fr

http://meghatropiques.ipsl.polytechnique.fr/available-documents/mttm/index.html