P13R.12

AUTOMATIC ESTIMATION OF RAINFALL FIELDS FOR HYDROLOGICAL APPLICATIONS: BLENDING RADAR AND RAIN GAUGE DATA IN REAL TIME

Carlos Velasco-Forero[1,*], Daniel Sempere-Torres[1], Rafael Sánchez-Diezma[1],
Eduardo Cassiraga[2], Jaime Gómez-Hernández[2]

(1) Grup de Recerca Aplicada en Hidrometeorologia. Universitat Politècnica de Catalunya, Barcelona (Spain)
(2) Departamento de Ing. Hidráulica y Medio Ambiente. Universitat Politècnica de Valencia, Valencia (Spain)

1. INTRODUCTION

Rainfall is one of the most important inputs to hydrological models, and rain gauge and weather radar are probably the two most widely sensors used in rainfall measurement. Unfortunately, operational rain gauge networks are not usually able to fulfill the requirements for hydrological modeling. Low spatial density and the irregular location of their measurement sites typically don't allow capture the large variability in space and time of precipitation fields. In this sense, weather radar could help to improve the knowledge of rainfall fields because radars provide an indirect measurement (reflectivity) of rainfall fields but with high resolution both in time (5 - 10 min) and space (1 km^2). These complementary characteristics are the motivation to work with both types of information. Therefore, an objective from the initial studies of hydrological applications of meteorological radar has been the development of methodologies to estimate rainfall fields merging radar and rain gauge data. In this sense several attempts have been reported previously, starting with the simplest formulation, finding constant multiplicative calibration factors (Wilson and Brandes 1979), and continuing through statistical approaches based on multivariate analysis (Hevesi et al. 1992) on radar rain gauge distribution probability analysis (Calheiros and Zawadzki 1987, Rosenfeld et al. 1994), on geostatistical estimators (Krajewski 1987, Creutin et al. 1988, Azimi-Zonooz et al. 1989, Seo 1998, Sinclair and Pegram 2005), or on Bayesian methods (Todini 2001).

Main problem for an implementation in real-time of geostatistical techniques is the definition of a valid spatial variability models (correlograms or variograms) from observations. In these conditions, the definition of valid spatial variability models (e.g., correlograms) has not been solved in a satisfactory way to avoid manual analysis or a priori simplified assumptions.

In this paper, an automatic technique to compute rainfall fields blending radar and rain gauges, which is based on kriging with external drift, is described and its performance is evaluated. The main interest of this new methodology is that it can be fully implemented in a real-time framework because the correlograms are automatically computed at each time step applying a fast approach based on the FFT.

[*] *Corresponding author address*: Carlos VELASCO, GRAHI-UPC Gran Capità 2-4 Ed. Nexus D-102. E08034-Barcelona (Spain). e-mail: velasco@grahi.upc.edu

2. THE PROPOSED ALGORITHM

2.1 Automatic Definition of Valid Correlograms

All geostatistical estimators (kriging) require knowing or to define previously a spatial variability model (covariance, correlogram or semivariogram) that provides the spatial continuity of estimated fields, and guarantee the solution of its equations system. The traditional (parametric) modeling approach to find these spatial models considers solely positive linear combinations of covariance or semivariogram basic models known to be positive definite under some restrictive conditions. These basic structures are closed form analytical functions defined by a few parameters, typically a range and a relative contribution to the total variation, hence the term "parametric" model. This modeling method requires one to decide on the number and the types of the basic covariance structures, and then determine their model parameters (range and contribution). In practice, parameters of basic models that summarize the experimental covariance structure are most often chose by trial and error, based on particular preferences, or applying some optimization algorithm based on goodness-of-fit or cross-validation. However, the restriction to linear combinations could be sometimes rather unfair, and also determining the linear combination that bests fits experimental data could be difficult and high time consuming. The prior decision about number and types of component structures is rather subjective. Also, sometimes the fit between the analytical model and the experimental values may be poor because the limitation of each structure being one of the few classic positive definite basic models. Different fitted-models and diverse methodologies to define these valid spatial variability models of rainfall fields have been reported in previous works, but each one has defined different spatial models according to the best adjustment with the specific data used and actually none have become a reference procedure. Exhaustive description of adjustment techniques and main difficulties on identification and calibration of spatial variability models of rainfall using the parametric modeling approach are reported in Bacchi and Kottegoda (1995).

As an alternative to the traditional methodologies, Yao and Journel (1998) have proposed a non-parametric methodology to find definite positive spatial covariance maps. This technique looking for covariance values that are not related by any analytical formulae, but provide a good fit of the corresponding experimental covariance. Principal advantage of this nonparametric technique is that not call for any prior choice of an analytical model, frees the modeling of coregionalization model, and guarantee valid and unique solution of kriging system. This nonparametric covariance modeling technique consists of transforming the experimental covariance values into spectral density maps in the frequency domain using FFT. Following Bochner's theorem (Bochner 1949), the spectral density maps are smoothed in this frequency domain under positivity constraints. Then, the FFT inverse transform of the smoothed spectral density tables provides positive definite covariance tables. This technique can be applied to define cross covariance models and to find valid variability models on one, two, or three dimensions. Therefore, this technique does not call for any previous choice of an analytical model, the transformation between spatial and frequency domain is done vary fast using FFT, and the constraints for positive definiteness in the frequency domain are easy to implement. Final valid covariance values are thus built from a FFT "roundtrip" with multiple intermediate smoothing processes. Full description of this algorithm could be found in Yao (1998).

Non-parametric techniques have not been applied previously in the context of radar – rain gauge data combination to estimate rainfall fields. In this paper, we propose use the non-parametric technique described above to define quickly and automatically bi-dimensional valid correlograms from radar data. This methodology allows compute different and valid spatial variability model for each time step in operational frameworks of merging radar – rain gauge process.

2.2 Kriging with an external drift

Our objective is to estimate rainfall fields that fit locally with direct (but scarce) rain gauge measurements and spatially with the shapes of exhaustive (but indirect) radar fields in framework of real-time combination of radar and rain gauge data by some kriging technique.

If exhaustive auxiliary information is available and correlated with the target variable, kriging with an external drift (KED) should be used in estimation process. KED assumes that the estimations should be modeled as a drift term plus a residual. The drift term is an unknown linear function of some auxiliary variables that is defined externally. The application of KED requires modeling the covariance of the residuals. Full description of KED equations can be seeing in literature (Wackernagel 1995, Goovaerts 1997, Hengl et al. 2003).

Assuming that we want to impose the spatial pattern of the radar data to the interpolated rainfall field and that the relation between rain gauge and radar data is linear, we propose the follow methodology in order to obtain rainfall fields by KED:

1. Calculate a drift map from the rainfall radar field by ordinary kriging for whole estimation area. This drift field is computed using only the rainfall radar data collocated on rain gauge locations. The valid correlogram map for this initial estimation is computed from the complete radar rainfall field using the nonparametric modeling technique. Different algorithms that the proposed here could be employed to compute the drift field as well (e.g., by window average). However posterior use of this drift map to estimate values conditioned by predefined locations induces to take account these positions into its definition process as it is proposed here.
2. Obtain the residual map subtracting the drift map calculated in step 1 from the radar field.
3. Apply the nonparametric modeling technique to the residual map in order to calculate a valid residual correlogram map.
4. With a valid correlogram map from residuals, KED algorithm can be applied to estimate rainfall fields using rain gauge data as primary variable, radar data as secondary variable, and that residual correlogram map as spatial variability model.

3. RADAR AND RAIN GAUGES DATA

Radar data used in this study were measured with the C-band radar of the Spanish Meteorological Institute (INM) located in Corbera de Llobregat (near Barcelona, Spain). Domain study was defined as a square region of 140 x 140 km around this radar. Inside of this domain, data of 77 rain gauges of Spanish SAIH network were selected as well (Figure 1).

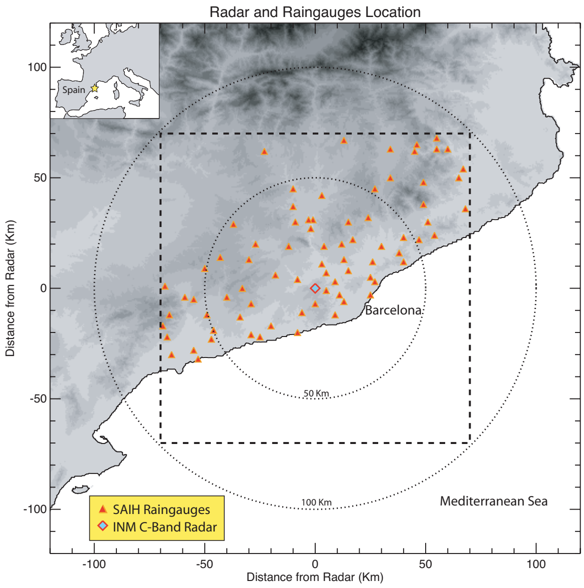

Figure 1. Map of the Barcelona area, showing the Barcelona C-band radar location (diamond), rain gauges (triangles) and range rings at 50 and 100 km from radar. Dashed square delimits the 140 x 140 km region that was used as study area.

A database of 275 hourly radar rainfall images corresponding to six different rainfall events was used to estimate rainfall fields blending them with hourly data of the rain gauges. Main characteristics of these case events are summarized in Table 1. Before merging radar fields with rain gauges data, some pre-processing techniques were applied to radar data sets. First, ground clutter echoes and partial screening effects were detected and corrected (Sánchez-Diezma et al. 2001, Sempere-Torres et al.

2002). Then, Z-R relationship proposed by Marshall and Palmer (1948) was applied to compute the intensity values associated with the levels of reflectivity of first PPI of radar fields. Finally, accumulated rainfall radar images were computed using an algorithm (Bellon et al. 1991) where the precipitation field is assumed to move at constant velocity and to vary linearly in intensity with time between each time interval.

TABLE 1. Main Characteristics of Case Studies.

Date	Duration (h)	Max. Accum. Rainfall (mm)
10th June 2000	23	223.8
28th September 2000	18	39.7
19th July 2001	25	38.6
15th November 2001	96	125.4
08th October 2002	38	193.8
30th March 2004	72	96.9

Adjustment between radar fields and rain gauges data for all time steps is presented in Figure 2. Radar values correspond with the collocated data on rain gauges locations. This figure clearly shows that radar underestimates rainfall values compared with rain gauges measurements.

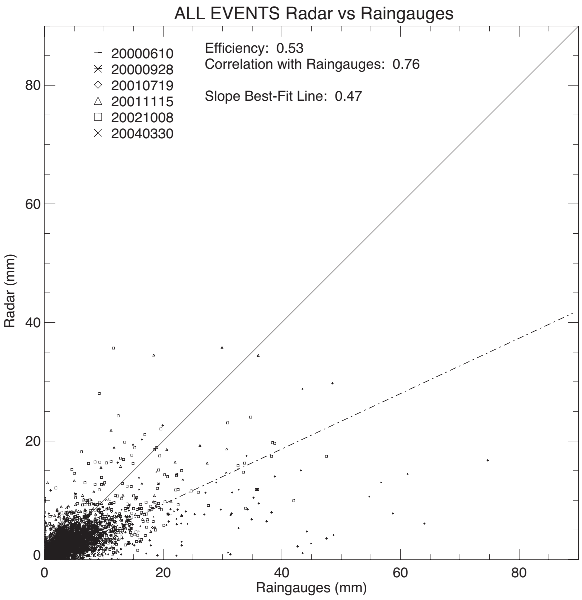

Figure 2. Scatter-plot between observed (rain gauges) and radar rainfall data. Data of all time steps of the six case studies are shown. Solid line shows the perfect agreement line, and dashed line shows the best-fit line of data.

4. RESULTS

We are looking for estimated rainfall fields with quantitative characteristics of rain gauge measurements but with spatial tendencies observed in radar field. Therefore to analyze the performance of KED estimator to assent this objective, two different processes were employed.

First, cross-validation technique was used to evaluate quantitative performance of KED estimator about to estimate with accuracy the rain gauges observations. In this technique, the primary data locations are systematically suppressed one at a time (one known rain gauge measurement) and the value at that location is predicted using only the remaining data locations via KED. Then the cross validation error, defined as the difference between estimated and known (but removed) value, is computed. Cross-validation estimations were independently calculated for each case study and for each time step.

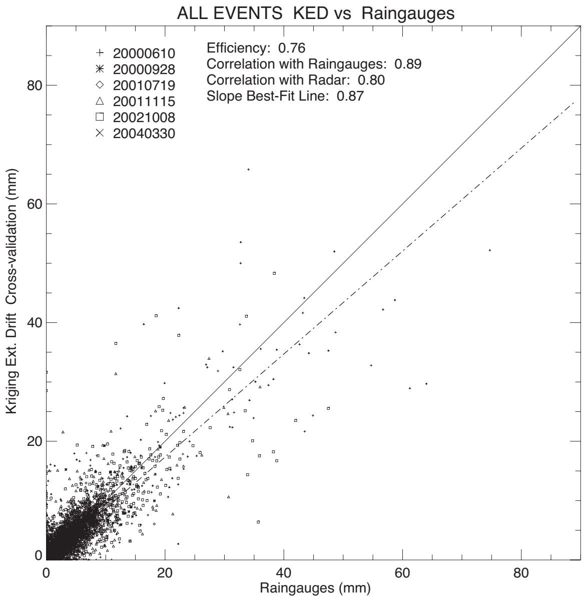

Figure 3. Scatter-plot between observed (rain gauges) and estimated rainfall from KED. Solid line shows the perfect agreement line, and dashed line shows the best-fit line of data

Figure 3 shows the adjustment between cross-validation estimations with the observed rain gauges data for all time steps. These KED estimations show a better agreement with rain gauges observations than original radar values (shown in Figure 2). Nash-Sutcliffe's efficiency of cross-validation estimations is equal to 0.76 instead of 0.53 for radar data, showing a best fit with observed rainfall values on rain gauges. Also, correlation between cross-validation estimations and rain gauges data is better than that computed with radar data (0.89 instead of 0.76). Finally, a reduction of the underestimation tendency of original radar data is evident on KED estimations. Slope of best-fit line between KED estimations and rain gauges (0.87) is more close to one than slope of radar data (0.47).

Additionally, temporal evolution of mean logarithmic bias of radar data and cross-validation estimations was analyzed and compared to illustrate the temporal performance of KED estimator. Mean field rainfall logarithmic bias, is defined as:

$$\beta_t = \frac{1}{n}\sum_{i=1}^{n}\log_{10}\left(\frac{G_{i,t}}{R_{i,t}}\right) \qquad (1)$$

where $G_{i,t}$ is hourly rain gauge rainfall (mm) at gauge i for hour t, $R_{i,t}$ is the radar rainfall (or cross-validation

KED estimation) (mm) at gauge *i* for hour *t*, and *n* is the number of radar-gauge pairs data available for that hour. Logarithmic mean biases computed for each time step from cross-validation estimations are compared with those computed from observed radar data in Figure 5. A bias of –1 indicates that there are less than 10 radar-gauge pairs data available for that hour. From this figure, it can be seen that the scatter of the estimated bias was less than that of the observed, and its values were closer to 0 than observed bias for all events. These results seem indicate KED estimator have potentiality to compute values of rainfall close with the observations of some hypothetical rain gauge at same point.

Finally, rainfall fields were estimated on all study area using KED, and then were compared with original radar field values. As estimated rainfall fields would preserve only general tendencies but not the particular values of radar fields, comparisons between rainfall fields were made using spectral analysis. The comparison skill used is the slope of the best linear fit to the logarithm of the power spectrum of the rainfall field. The power spectrum derived from the fields is supposed to follow a power-law in frequency (see e.g. Pegram and Clothier 2001). If the power Fourier spectrums of two different rainfall fields have similar slopes, their spatial tendencies would be similar. Therefore, Figure 4 shows the agreement between the slopes of spectrum of rainfall radar fields and rainfall fields estimated by KED at each time step. Time steps with less than 10 radar-gauges pairs data available for that hour were not included in the figure. From this figure, it can be seen that the KED-estimated rainfall fields have values of slopes close with those obtained of the radar rainfall fields in the most part of compared time steps. Dispersion of points is reduced as well.

From these results, we can conclude that KED estimator with the automatic (nonparametric) definition of correlograms could estimate rainfall fields based in rain gauge and radar data with a good punctual agreements with rain gauges and with similar spatial patters of those observed in the radar rainfall fields.

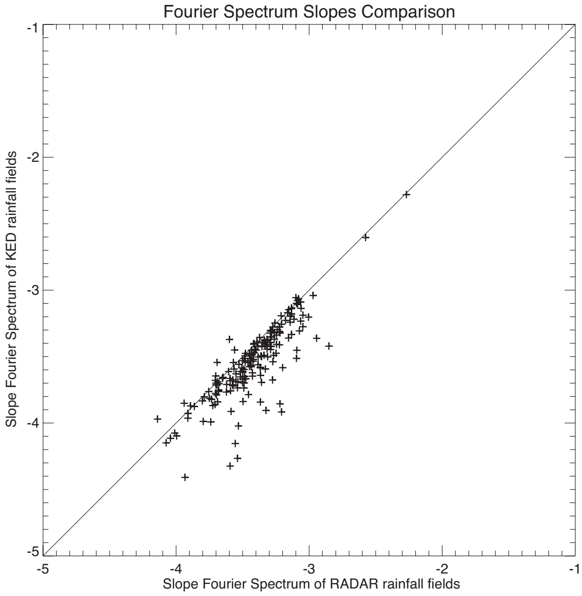

Figure 4. Comparison between slopes of Fourier Spectrum of radar rainfall fields and KED rainfall fields. Each point corresponds with a time step with more than 10 radar-gauge pairs data available for that hour. Solid line shows the perfect agreement line.

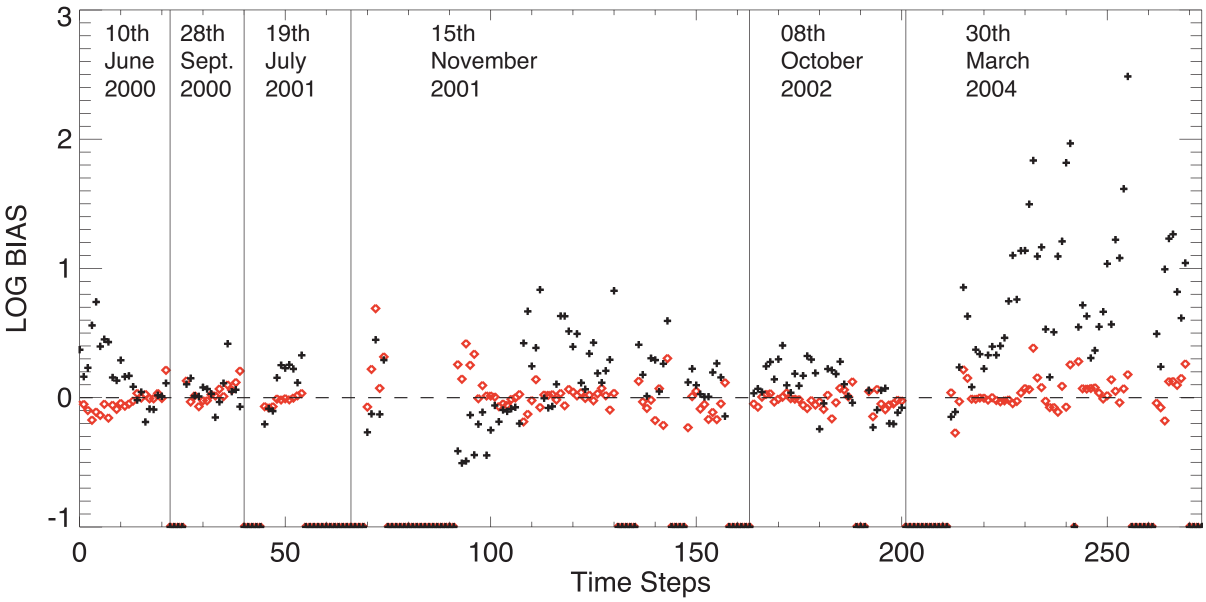

Figure 5. Comparison of the observed (black cross) and the cross-validation KED estimated (red diamond) mean field logarithmic bias of the events. A bias value of –1 indicated that there are less than 10 radar-gauge pairs data available for that hour.

5. CONCLUSIONS

This paper has examined the use of kriging with external drift (KED) methodology for estimate rainfall field using radar and rain gauge data. Spatial variability models (correlogram) for this estimator were not defined in traditional way, but estimated automatically using a newly technique based on a nonparametric FFT-smoothing process.

This automatic technique was used to define a different valid 2D correlogram for each time step analyzed. These correlograms guaranteed, in all cases, the solution of KED equations systems.

Our goal is to estimate rainfall fields with quantitative characteristics of rain gauge measurements but with spatial tendencies observed in radar field. Therefore, cross-validation, temporal logarithmic mean bias, and spectral analysis were applied to the KED estimations. Analysis of results seem indicate that KED technique with the automatic definition of correlograms could estimate based on rain gauge and radar data rainfall fields with a good punctual agreement with rain gauges and with similar spatial patters of those observed in the radar rainfall fields.

This automatic technique to estimate rainfall fields is extremely promising, particularly in time real applications of radar and rain gauge data for operational purposes.

Acknowledgements: This project has been carried out in the framework of the EC projects VOLTAIRE (EVK2-CT-2002-00155) and FLOODSITE (GOCE-CT-2004-505420). Spanish Institute of Meteorology (INM) provided the radar data. First author also is thankful to the Polytechnic University of Catalonia (UPC) for funding his doctoral studies through the scholarship “UPC per a la Formació de Personal Investigador”.

5. *References*

Azimi-Zonooz, A., W. F. Krajewski, D. S. Bowles, and D.-J. Seo. 1989. Spatial rainfall estimation by linear and non-linear co-kriging of radar-rainfall and raingage data. Stochastic Hidrol. Hydraul. 3:51-67.

Bacchi, B., and N. T. Kottegoda. 1995. Identification and calibration of spatial correlation patterns of rainfall. Journal of Hydrology 165:311-348.

Bellon, A., F. Fabry, and G. L. Austin. 1991. Errors due to space/time sampling strategies in high resolution radar data used in hydrology. Pages 840-843 in 25th Radar Meteor. Conf. AMS, Paris.

Bochner, S. 1949. Fourier Transforms. Princeton University Press, London.

Calheiros, R. V., and I. Zawadzki. 1987. Reflectivity-rain rate relationship for radar hydrology in Brazil. J. Climate Appl. Meteorol. 26:118-132.

Creutin, J. D., G. Delrieu, and T. Lebel. 1988. Rain measurement by raingage-radar combination: a geostatistical approach. J. Atmos. Oceanic Technol. 5:102-115.

Hevesi, J. A., J. D. Istok, and A. L. Flint. 1992. Precipitation Estimation in Mountanious Terrain using Multivariate Geostatistics. Part I: Structural Analysis. Journal of Applied Meteorology 31:661 - 676.

Krajewski, W. F. 1987. Cokriging radar-rainfall and rain gage data. J. Geophys. Res. 92:9571-9580.

Marshall, J. S., and W. K. Palmer. 1948. The distribution of raindrops with size. Journal of Meteorology 5:165-166.

Pegram, G. G. S., and A. N. Clothier. 2001. High resolution space-time modelling of rainfall: the "String of Beads" model. Journal of Hydrology 241:26-41.

Rosenfeld, D., D. Atlas, D. B. Wolff, and E. Amitai. 1994. The window probability matching method for rainfall measurements with radar. J. Appl. Meteor. 34:682-693.

Sánchez-Diezma, R., D. Sempere-Torres, G. Delrieu, and I. Zawadzki. 2001. An Improved Methodology for ground clutter substitution based on a pre-classification of precipitation types. Pages 271-273 in 30th International Conference on Radar Meteorology, Munich, Germany.

Sempere-Torres, D., R. Sánchez-Diezma, M. Berenguer, R. Pascual, and I. Zawadzki. 2002. ASCMORE: an algorithm to control radar rainfall measurement stability using mountain returns in real time. Geophysical Research Abstracts 4:6319.

Seo, D.-J. 1998. Real-time estimation of rainfall fields using radar rainfall and rain gage data. Journal of Hydrology 208:37-52.

Sinclair, S., and G. Pegram. 2005. Combining radar and rain gauge rainfall estimates using conditional merging. Atmospheric Science Letters (In press).

Todini, E. 2001. A Bayesian technique for conditioning radar precipitation estimates to rain-gauge measurements. Hydrology and Earth System Sciences 5:187 - 199.

Wilson, J. W., and E. A. Brandes. 1979. Radar measurement of rainfall- a summary. Bull. Amer. Meteor. Soc. 60:1048-1058.

Yao, T. 1998. Automatic Covariance Modeling and Conditional Spectral Simulation with Fast Fourier Transform. Doctoral. Stanford University, Stanford, California.

Yao, T., and A. G. Journel. 1998. Automatic Modeling of (Cross) Covariance Tables Using Fast Fourier Transform. Mathematical Geology 30:589 - 615.