Spatial - temporal modeling

Author(s):
Mussio, Luigi; Nocera, Rossella; Poli, Daniela

Publication Date:
1999

Permanent Link:

https://doi.org/10.3929/ethz-a-004328842 →

Spatial - Temporal Modeling

Luigi Mussio
DIIAR - Politecnico di Milano
luigi@ipmtf2.topo.polimi.it

Rossella Nocera
DIEMA - Università degli Studi di Reggio Calabria

Daniela Poli
DIIAR - Politecnico di Milano

Abstract
The basic idea underlying the proposed approach to spatial-temporal modeling is to combine, at the different levels, the potential of various techniques and to exploit their performance. A three level procedure is used. Firstly low level polynomial interpolation removes a general trend; then finite element method follows to match local variations. The last step is performed assuming the topography to be a realization of a continuos, normal, stationary, ergodic and isotropic (or orthogonal decomposed) stochastic process. Thus a filtering technique, based on covariance estimation and collocation, is applied. Finally the three levels can be computed together by using an integrated approach.

Integrated geodesy approach

The aim of geodesy and geomatics is the determination of the position of points, together with their variation in time and the representation of them, in various suitable forms.
The geometric observables in geodesy and related sciences which can be determined by satellite geodesy, surveying, photogrammetry, space photogrammetry and remote sensing, are all simple, or more complex functions of the relative position of two points. They are influenced by various physical fields, like the earth's gravitational field and the field of atmospheric refraction. Geodetic networks, in their various forms, like leveling networks, triangulation and trilateration networks or photogrammetric blocks, allow for the transition from geometric observables to the determination of the (relative) position of the network points. When measurements are repeated in time, variations in (relative) position can be determined. The imposition of an arbitrarily chosen reference system is needed to remove the indetermination of origin, orientation and scale.
The relation between the observables and the geometric parameters, which describe the position of the network points, is expressed in functional models. These are often non-linear; therefore have to be linearized (using approximate values for the parameters) to provide a linear system of observation equations. The rank defect of this system is removed by the introduction of a reference system. The relation between the geodetic observables and the physical field(s) can conveniently be expressed in stochastic models. These models, linearized and interpreted stochastically, provide covariance functions (invariant with respect to suitable groups of parameter transformations) with which one can obtain the stochastic signal present in the observables. Stochastic models may also concern geometric parameters, like systematic or pseudo-systematic measurement errors, heights, variations of height or position, map projections, the gray level or the color of images, the contour and other features of figures, the surface and other features of objects. In integrated geodesy (in its broadest sense), after the formulation of the functional and stochastic models which tie the observables to the stochastic and not-stochastic parameters, the complete set of equations and unknowns is solved simultaneously in one single adjustment. The computation of the solution is nowadays feasible. However, because of its complexity it has to be regarded as a final analysis, following both the traditional network adjustment and the digital modeling of the physical fields and/or the geometric parameters treated like fields.

Before dealing with the separate topics, the formal structure of observation equations is considered again with more details.

The geometric observables are generally non-linear functions of the difference of the point-positions. These could depend on time; moreover other terms are involved like gravity field effects, atmospheric refraction influences and calibration parameters, according to the data acquisition procedures.
The eventual movements of point-positions can be modeled with suitable functions of time. The parameters of the movement laws may have a similar behavior in neighboring points, so that they can be modeled by suitable functions of the point-positions. In such a way, not only the position parameters, but also the movements (displacements and deformations) can be related to the geometrical observations. Notice that dense fields of points can be modeled as a function of geographical positions in order to obtain continuos digital models of morphological features of figures and morphological features and shapes of objects.
Apart from their eventual time dependence, the equations of the geometric observables are represented by relatively simple and well known geometrical functions. The time dependence, if any, and the morphological aspects can be obtained by data fitting with polynomials or spline functions and, in the frequency domain, by Fourier or wavelets analysis. The approach of stochastic processes is often preferred when the time series is long or the field of points is large. In these cases, covariance functions are defined a priori with respect to the parameters of a coordinate transformation group. Notice that the number of repetitions is usually large in networks specifically designed for the monitoring of critical deformations, where the number of network points is often relatively small, as well as for the modeling of morphological aspects. On the contrary, for the monitoring of slowly moving points the number of network points is usually large, but the number of repetitions, if any, is often relatively small. Long time series and large fields of points are generally rare.

Considering a possible use of the collocation method, it is obvious that time invariance can be claimed only for long time series, whilst space invariance plus spherical isotropy (or orthogonal decomposition) can be claimed only for large fields of points.

In the traditional approach different systems are solved successively. Firstly field influences and calibration parameters are determined, then for each epoch a network adjustment is carried out, after which data fitting in applied to describe eventual point movements as a function of time. Finally characteristic parameters and/or morphological aspects are modeled as functions of geographical position. In the integrated geodesy approach all systems are solved simultaneously. Thus after the linearization of the observation and pseudo-observation equations, the observables and the other data are collected in a unique system containing uncorrelated unknown parameters, as well as correlated ones, that can be interpreted as stochastic signal to filter from the random noise.
Notice that it is necessary to perform preceding separate adjustments. Indeed the covariance matrix of the signal is obtained from estimates for the unknown parameters or residuals; moreover the variance of the noise is assumed equal to the sigma

naught square obtained in preceding separate adjustments. This can give some trouble in the fixing of the weights of the different elements. However by repeating the integrated geodesy approach adjustment, the uncertainty about the weight ratios can be eliminated and suitable values for the weights can be established, finding a reproduction point of them.
Moreover all the data are supposed outlier free. However because outliers occur in the data, due to gross errors and/or unmodeled effects, a suitable strategy combining robustness and efficiency has to be used. Indeed robust procedures are useful for the identification of suspected outliers, whilst the least squares are very powerful for testing about acceptance or rejection.
The rank defect of the set of matrices, which make up the unique system of equations, is easily determined. Indeed a system containing repeated network observation has rank defect equal to the number of indeterminable parameters of the approximation functions which describe variation in time of the point-positions. Moreover the system containing those observations, together with some continuous digital models, has generally same rank defect.

Approximation theory [De Haan et al., 1994]

A unique system collects all the observation and pseudo-observation equations, referred to observables and other data y, and contains uncorrelated unknown parameters x , as well as correlated ones that can be interpreted as stochastic signal s to filter from the random noise n:

$$y = Ax + Bs$$

The use of both stochastic and non-stochastic parameters causes the need to introduce a hybrid norm:

$$\frac{1}{2}\begin{bmatrix}\hat{s}^t & \hat{n}^t\end{bmatrix}\begin{bmatrix}C_{ss}^{-1} & 0\\ 0 & P/s_n^2\end{bmatrix}\begin{bmatrix}\hat{s}\\ \hat{n}\end{bmatrix} + l^t\left(A\hat{x} + B\hat{s} - \hat{n} - y_0\right) = min$$

where y_0 indicates the observations, $\hat{x}$, $\hat{s}$, $\hat{n}$ the estimated values of x, s and n respectively, C_{ss} the covariance matrix the signal, s_n^2 the variance of the noise, P the weight matrix of the observations and l a vector of Lagrange multipliers.
Unfortunately the solution of this kind of system isn't practically computable. Indeed it presents heavy computations, from the point of view of both storage requirements and time consuming, because the unknown parameters are split in two separate parts. Therefore the above mentioned system of observation equations is rewritten as:

$$y = Bs$$

with s containing both stochastic and non-stochastic parameters: $s^t = \begin{bmatrix}x^t & s^t\end{bmatrix}$, and the design matrix B defined as: $B = \begin{bmatrix}A & B\end{bmatrix}$, expressing both to chosen functional and stochastic modeling. The observed quantities y_0 are related to the estimates $\hat{s}$ of s by the same linearized model:

$$B\hat{s} - \hat{n} - y_0 = 0$$

The covariance matrix C_{ss} for the newly defined signal s contains four blocks, two diagonal blocks containing the covariance matrices of the stochastic and non-stochastic parts of the signal and two zero off-diagonal blocks:

$$C_{ss} = \begin{bmatrix}hI & 0\\ 0 & C_{ss}\end{bmatrix}$$

The covariance matrix of the stochastic parameters is determined by one or more auto and crossvariance functions which can be estimated empirically with the results of preceding separate adjustments. The covariance matrix of the non-stochastic parameters is a diagonal matrix, the elements of which (if not related constraints) have to be chosen in balance with the covariances of the stochastic parameters. In such a way, the solution is not constrained too much to either type of parameters.
The general variance of the noise which also has to be known a priori can be assumed equal to the estimated variance factor of preceding separate adjustments.
The generalized least squares criterion can be used to minimize contemporaneously the norm $\hat{s}^t C_{ss}\hat{s}$ and the norm of the residuals of the observation equations $\hat{n}^t P\hat{n}/s_n^2$:

$$\frac{1}{2}\begin{bmatrix}\hat{s}^t & \hat{n}^t\end{bmatrix}\begin{bmatrix}C_{ss}^{-1} & 0\\ 0 & P/s_n^2\end{bmatrix}\begin{bmatrix}\hat{s}\\ \hat{n}\end{bmatrix} + l^t\left(B\hat{s} - \hat{n} - y_0\right) = min$$

According to this criterion, the estimate for the signal and noise become:

$$\hat{s} = C_{ss}B^t\left(BC_{ss}B^t + s_n^2P^{-1}\right)^{-1}y_0$$

$$\hat{n} = s_n^2P^{-1}\left(BC_{ss}B^t + s_n^2P^{-1}\right)^{-1}y_0 = y_0 - B\hat{s}$$

and the covariance propagation law permits to compute their covariance matrices:

$$C_{ee} = C_{ss} - C_{\hat{s}\hat{s}} = C_{ss}\left(I - B^t\left(BC_{ss}B^t + \sigma_n^2P^{-1}\right)^{-1}BC_{ss}\right)$$

$$C_{nn} = s_n^4P^{-1}\left(BC_{ss}B^t + s_n^2P^{-1}\right)^{-1}$$

being: $e = s - \hat{s}$ the estimation error of the signal.
The computation of the preceding expressions requires the solution of a system with dimension m, equal the number of observations. However it would be more convenient to have analogous expressions, which require the solution of a system with dimension $n < m$, equal to the number of parameters. A further requirement would be the absence of inverse matrices which contain inverse matrices. Both can be achieved by the application of the two famous theorems of linear algebra which are stated below:

$$\left(Q \pm RST\right)^{-1} = Q^{-1} \mp Q^{-1}R\left(S^{-1} \pm TQ^{-1}R\right)^{-1}TQ^{-1}$$

$$Q^{-1}\left(Q^{-1} \pm S\right)^{-1}Q^{-1} = \left(Q \pm QSQ\right)^{-1}$$

Therefore the estimate for the noise can be rewritten as:

$$\hat{n} = y_0 - B\left(\left(B^tPB\right)^{-1} - s_n^2\left(B^tPBC_{ss}B^tPB + s_n^2B^tPB\right)^{-1}\right)$$
$$B^tPy_0 = y_0 - B\hat{s}$$

Furthermore taking into account the last expression, the estimate for the signal becomes:

$$\hat{s} = \left(B^tPB\right)^{-1}B^tPy_0 - s_n^2\left(B^tPBC_{ss}B^tPB + s_n^2B^tPB\right)^{-1}B^tPy_0$$

With these new expressions, the covariance propagation law permits the computation of the corresponding covariance matrices, in equally convenient forms. Thus the covariance matrices of the estimation error of the signal and the residual noise become, respectively:

$$C_{ee} = C_{ss} - C_{\hat{s}\hat{s}} = \sigma_n^2 \left(B^t PB\right)^{-1} + \\ -\sigma_n^4 \left(B^t PBC_{ss} B^t PB + \sigma_n^2 B^t PB\right)^{-1}$$

$$C_{\hat{n}\hat{n}} = \sigma_n^2 P^{-1} - BC_{ee} B^t = \sigma_n^2 \left(P^{-1} - B\left(B^t PB\right)^{-1} B^t \right) + \\ + \sigma_n^4 B\left(B^t PBC_{ss} B^t PB + \sigma_n^2 B^t PB\right)^{-1} B^t$$

Notice that, as already said before, heavy computation doesn't occur, from the point of view of both storage requirements and time consuming.

Let recall that, in case both the matrices B and P are identity matrices, the problem becomes a filtering of the stochastic signal from the random noise only. Therefore:

$$\hat{s} = C_{ss}(C_{ss} + s_n^2 I)^{-1} y_0 = y_0 - \hat{n}$$
$$\hat{n} = s_n^2 (C_{ss} + s_n^2 I)^{-1} y_0 = y_0 - \hat{s}$$

and the covariance matrices of the estimation error of the signal and the residual noise become, respectively:

$$C_{ee} = C_{ss} - C_{\hat{s}\hat{s}} = C_{ss} - C_{ss}(C_{ss} + \sigma_n^2 I)^{-1} C_{ss} = \sigma_n^2 I - C_{\hat{n}\hat{n}}$$
$$C_{\hat{n}\hat{n}} = \sigma_n^4 (C_{ss} + \sigma_n^2 I)^{-1}$$

In this case, the calculation goes quickly, in a very easy way.

Finally by replacing $\hat{s}$ with $\hat{x}$ and B with A, forcing a little the previous general expressions and evaluating s_n^2 is equal to zero, the classical least squares solution is achieved, because second term of both expressions vanishes:

$$\hat{x} = \left(A^t PA\right)^{-1} A^t Py_0$$
$$\hat{v} = y_0 - A\hat{x} = y_0 - \hat{y}$$

Regarding the covariance matrices, defining the cofactor matrices without sigma naught, as usual in the standard form, one gets:

$$C_{\hat{x}\hat{x}} = s_0^2 \left(A^t PA\right)^{-1}$$

$$C_{\hat{v}\hat{v}} = s_0^2 \left(P^{-1} - A\left(A^t PA\right)^{-1} A^t\right) = s_0^2 P^{-1} - AC_{\hat{x}\hat{x}} A^t = \\ = s_0^2 P^{-1} - C_{\hat{y}\hat{y}}$$

where the second term of both expressions vanishes once more.

The generalized least squares

The generalized least squares criterion, expressed in the formulation of the collocation method, can provide, besides an estimate for a filtered signal, also an estimate for a predicted signal. Indeed the stochastic parameters can also be estimated in points which don't make part of the data. One has to keep in mind however, that only the properly called stochastic parameters can be estimated.

Consequently the covariance matrix consists of the properly called stochastic parameters only and the crosscovariance matrix between the filtered and the predicted signal is divided in two parts: one containing the covariance between the predicted signal and the properly called stochastic parameters in the filtered signal and one identically zero. This null matrix is exactly the reason of the impossibility to predict the parameters which are strictly non stochastic.
Given the functional:

$$\frac{1}{2}\left[\hat{s}^t \hat{t}^t \hat{n}^t\right] \begin{bmatrix} W^{-1} & 0 \\ 0 & P/s_n^2 \end{bmatrix} \begin{bmatrix} \hat{s} \\ \hat{t} \\ \hat{n} \end{bmatrix} + l^t \left(B\hat{s} + 0\hat{t} + \hat{n} - y_0\right) = min$$

$$\text{where: } W = \begin{bmatrix} C_{ss} & C_{st} \\ C_{ts} & C_{tt} \end{bmatrix}$$

being l a vector of Lagrange multipliers, and taking into account the expression of the estimate for the filtered signal one has:

$$\hat{t} = C_{ts} B^t \left(BC_{ss} B^t + s_n^2 P^{-1}\right)^{-1} y_0 = \\ = C_{ts} B^t PB\left(B^t PBC_{ss} B^t PB + s_n^2 B^t PB\right)^{-1} B^t Py_0$$

or:

$$\hat{t} = C_{ts} z$$

with z a service vector:

$$z = B^t PB\left(B^t PBC_{ss} B^t PB + s_n^2 B^t PB\right)^{-1} B^t Py_0$$

which is to be computed once at the end of the filtering (notice: $z = \left(C_{ss} + \sigma_n^2 I\right)^{-1} y_0$ if $B = P = I$). Applying the covariance propagation law, the covariance matrix of estimation error of the predicted signal becomes:

$$C_{ee} = C_{tt} - C_{\hat{t}\hat{t}} = \\ = C_{tt} - C_{ts} B^t PB\left(B^t PBC_{ss} B^t PB + \sigma_n^2 B^t PB\right)^{-1} B^t PBC_{st}$$

Unfortunately this expression isn't very convenient in computation and it is impossible to find others more suitable. Therefore its computation is usually omitted.
The variance of the noise can also be estimated a posteriori. Imposing its estimate to be unbiased, one obtains:

$$k = m - n + Tr\left(s_n^2 P^{\frac{1}{2}} B\left(B^t PBC_{ss} B^t PB + s_n^2 B^t PB\right)^{-1} B^t P^{\frac{1}{2}} \right)$$

where m is the number of observations and n the number of parameters (notice: $k = m - n$ if $\sigma_n^2 = 0$). Therefore the a posteriori estimate of the variance of the noise becomes:

$$\hat{s_n^2} = \hat{n}^t P\hat{n} / k$$

The same formula can be used for the a posteriori estimate of variances and therefore also of weights of a priori defined groups of observations.

With respect to the computability some considerations are made concerning the applications of the above mentioned theorems. As

already said before, a suitable application of these theorems provides systems of dimension $n < m$, without inverse matrices which contain inverse matrices. The formulae of the estimate of the filtered signal and the covariance matrix of its estimation error contain the expressions:

$$\left(B^t PB\right)^{-1} BPy_0 \quad ; \qquad \left(B^t PB\right)^{-1}$$

The solution of this system and the computation of the inverse matrix are standard procedures in any least squares problem and, if the design matrix B is of a network, are computable with direct solution algorithms which are capable to work with sparse matrices. The same formulae also contain in the expressions:

$$\left(B^t PBC_{ss} B^t PB + s_n^{\,2} B^t PB\right)^{-1} B^t Py_0 \quad ;$$
$$\left(B^t PBC_{ss} B^t PB + s_n^{\,2} B^t PB\right)^{-1}$$

The normal matrix $B^t PB$ was already obtained before. The covariance matrix: $C_{ss} = C_{ss} * S_{ss}$, of the properly called stochastic parameters is a sparse matrix when constructed by multiplying, according to Hadamard, the proper covariance matrix C_{ss} by a suitable finite covariance matrix S_{ss}.. Its sparseness depends on the persistence of correlation of the finite covariance functions. Its dispersion however is influenced by the numbering of the data points. The product of three sparse matrices $\left(B^t PB\right) C_{ss} \left(B^t PB\right)$ is a sparse matrix itself. Therefore the solution of the corresponding system can be computed with appropriate algorithms.
Let recall that, regarding the solution strategies, three classes of data have to be considered: network structures, dense but irregular fields of points, raster data and grid parameters. They imply different numerical techniques: direct algorithms (for sparse matrices), iterative algorithms and special tools involving regularity. The last two techniques are the above mentioned appropriate algorithms and the choice depends on the data regularity.

Covariance estimation [Bellone et al., 1996]

Covariance estimation and covariance function modeling are very important, because collocation filtering and prediction require appropriate models to interpolate the empirical covariance functions of the signal.
An hypothesis has been made: the data can be seen as realizations of a continuous and normal stochastic process which

$$C\left(P_1, P_2\right) = C\left(\left\|P_1 - P_2\right\|\right).$$

Notice that spatial stochastic processes can be assumed isotropic, as well as with orthogonal separability among the coordinates. The former uses the euclidean distance which is invariant on a circle (a sphere, a hypersphere); the latter uses the Manhattan distance which is invariant on a square (a cube, a hypercube). Notice that L_1 norm implies orthogonal separability and offers a big gain from the numerical point of view.
With $X\left(P_i\right)$ the n observations at the different points, the estimates of the empirical auto-covariance function at the interval $D^{(l)}$ are calculated from:

$$g\left(D^{(l)}\right) = \frac{1}{n}\sum_{i=1}^{n} v_i \frac{1}{n_i^{(l)}} \sum_{j=1}^{n_i^{(l)}} v_j^{(l)}$$

where:

$$\left\{v_j^{(l)} : P_j \Rightarrow D^{(l-1)} < \left\|P_i - P_j\right\| \le D^{(l)}\right\}$$

and:

$$v_k = x_k - \bar{x} \qquad k = 1, \cdots, n$$

whilst the estimates of empirical cross-covariance function at the interval $D^{(l)}$ are computed in similar way.
When the data points aren't equally spaced, the optimal interval size has to be found for the covariance estimates. In fact, an interval which is too small will contain only few data points, whilst an interval which is too large will contain data points which are too much dispersed. In both cases the estimates are suppressed.
A criterion for the choice of the interval including the first auto-covariance zone is maximizing the first autocovariance estimate as follows:

$$r_{(l)} : g\left(D^{(l)}\right) = max\left(g\left(D^{(l)}\right)\right)$$

where:

$$g\left(D^{(l)}\right) = \frac{1}{n}\sum_{i=1}^{n} v_i \frac{1}{n_i^{(l)}} \sum_{j=1}^{n_i^{(l)}} v_j^{(l)}$$

and:

$$\left\{v_j^{(l)} : P_j \Rightarrow 0 < \left\|P_i - P_j\right\| \le D^{(l)}\right\}$$

In the case of vector quantities, the determination of the optimal interval spacing is obtained by minimizing the trace of the covariance matrix of the uncorrelated noise which is invariant with respect to rotations.
Furthermore by using the orthogonal separability among the coordinates, it is easy to obtain covariance estimates for each row and each column. The median of the row covariance estimates and the median of the column covariance estimates give the empirical auto-covariance functions for the rows and the columns. The covariance estimation continues then by multiplying the row covariance estimates by the column covariance estimates, in order to obtain the covariance function of the data. The same procedure has to be done for the best fit covariance function and the noise variance.
In case of 3D (or ND) stochastic processes, the orthogonal separability among the three (or more) coordinates goes to the estimation of covariances for each row, each column and each pile (etc. if necessary). Successively the medians give the empirical auto-covariance functions and the product among the three (or more) covariance components supplies the covariance function of the data.
After empirically having estimated points of the covariance functions one has to interpolate them by using classes of approximation functions which are able to simulate the behavior of a covariance function.
Because covariance estimation doesn't satisfy positive definite property automatically, this last must be achieved by modeling the covariance estimates with a suitable set of positive definite models.
The best fit of the auto-covariance function of the signal is then chosen among some available models, namely, when euclidean distance is used in 2D stochastic processes:

$$g\left(D\right) = a\, exp\left(-bD\right)$$
$$g\left(D\right) = a\, exp\left(-bD^2\right)$$
$$g\left(D\right) = a\, exp\left(-bD\right)\left(1 - cD^2\right)$$
$$g\left(D\right) = a\, exp\left(-bD^2\right)\left(1 - cD^2\right)$$
$$g\left(D\right) = a\, exp\left(-bD\right) J_0\left(cD\right)$$

$$g(D) = a\,exp(-bD^2)J_0(cD)$$
$$g(D) = 2a\,exp(-bD)J_1(cD)/(cD)$$
$$g(D) = 2a\,exp(-bD^2)J_1(cD)/(cD)$$

or alternatively, when Manhattan distance is used:

$$g(D) = a\,exp(-bD)$$
$$g(D) = a\,exp(-bD^2)$$
$$g(D) = a\,exp(-bD)(1-cD^2)$$
$$g(D) = a\,exp(-bD^2)(1-cD^2)$$
$$g(D) = a\,exp(-bD)\cos(cD)$$
$$g(D) = a\,exp(-bD^2)\cos(cD)$$
$$g(D) = a\,exp(-bD)\sin(cD)/(cD)$$
$$g(D) = a\,exp(-bD^2)\sin(cD)/(cD)$$

where the smoothness given by the coefficient b is in both cases very high.
This list has been built according to the properties of covariance function: positive power spectrum, i.e. positive Fourier transform, and Schwarz inequality for vector processes.
New covariance function can be created from old by applying following fundamental theorems:

- a linear combination with positive coefficients;
- a product;
- a convolution;

The same list is used to interpolate a cross-covariance function: it is not correct in principle, but is acceptable in practice, provided that cross-covariance estimates are low enough. Many different approaches exist for the choice of a suitable class of approximation functions to interpolate empirical covariance estimates. A possible choice for the class of approximating functions is based on the analysis of the trend (concave or convex) in a region near the origin, on the presence of possible zeros and, if there are any zeros, on relative maxima and minima. Whatever the type of model chosen for the interpolation of the covariance estimates, the interpolating function can be imposed to coincide with the first empirical covariance estimate.

The classes of approximation functions provide covariance functions which have zero function values in at most a countably infinite number of points and asymptotically tend to zero only at infinity. Consequently a full covariance matrix results for any set of points. This can lead to insurmountable computational problem, with respect to both storage requirements and time consuming.
A finite covariance function, i.e. a function which differs from zero only in a small part of its domain, provides a sparse covariance matrix. A sparse covariance matrix doesn't present insurmountable storage requirements, as only non-zero elements are stored (along with some service vectors). The time consuming is also much lower, if a suitable ordering of the unknowns can be found.
In the one dimensional case, a reordering can be found which clusters the non-zero elements near the diagonal; moreover a direct solution of the system can then be found using band algorithms. In the 2D, 3D (ND), cases the non-zero elements in the matrix are fairly dispersed and it may not be possible to find a convenient ordering. In that cases, iterative solution algorithms have to be used which work with sparse storage schemes.
Furthermore in all cases, if there is regularity within the data, Toeplitz matrices have to be considered and Kronecker decomposition have to be performed too, when the dimension of the problems is greater than one.

As above mentioned, a function $h(z)$ obtained by the convolution any function $f(x)$ with itself:

$$h(z) = \int_{-\infty}^{+\infty} f(x)\,f(z-x)\,dx$$

has the properties of a covariance function. Because the convolution of a finite with itself is a finite function, finite covariance functions can be obtained by subsequent convolutions.
Besides finite covariance functions are obtained by multiplying the best fit by positive definite finite functions which are given by convolution of a finite function with itself, respectively, in the first isotropic case, considering a 2D stochastic processes:

$$\gamma(\Delta) = 5\pi - 15\Delta^2\pi/2 + \left(5\Delta + 20\Delta^3/3 - 5\Delta^5/12\right)\sqrt{1-\Delta^2/4} + \left(15\Delta^2 - 10\right)arcsin\left(\sqrt{\Delta/2}\right)$$
with $D \le 2$

$$g(D) = 0 \quad \text{with} \quad D \ge 2$$

and in the second with orthogonal separability case:

$$g(D) = 16/15 - 4D^2/3 + 2D^3/3 - D^5/30$$
with $D \le 2$

$$g(D) = 0 \quad \text{with} \quad D \ge 2$$

In case of 3D (or ND) stochastic processes, a library of auto-covariance function models, especially when they are finite covariance functions, is difficult to construct. On the other hand, the orthogonal separability among the three (or more) coordinates allows to use three (or more) times the one-dimensional auto-covariance function models. The same is true for the finite covariance functions and the product among the three (or more) covariance components supplies positive definite models.
The extension to the space-time domain implies similar procedures. Notice that the covariance estimation in the space-time domain requires as additional hypotheses:

- the irrotational condition between the time and the space coordinates;
- the separability between the two types of coordinates in the empirical estimate, so that the space coordinates are free to run at any fixed time and the time is free to run in any fixed place.

This means that each covariance matrix must be split in two factors, according to the Hadamard product:

$$C(DP, Dt) = C(DP) * C(Dt)$$

Cluster analysis could help to collect homogeneous covariance estimates, i.e. (intermediate) data with a similar behavior.

Finally the noise variance is found as:

$$s_n^2 = s^2 - s_s^2 = s^2 - a$$

and the eventual noise covariance can be found with a similar formula.

The Taylor-Karman structure

Prediction is an operation which, as already said before, can be performed in a lot of applications. Therefore not only the signal, but also its functionals can be estimated and, among these,

differential operators to the signal are very important, because they are linear and supply information about shape and pattern.
In the follows, the covariance propagation of a 2D isotropic stochastic process is investigated, showing the Taylor-Karman structure, when differential operators are applied.
The matrices of the two first derivatives of a function are respectively the so called Jacobian and Hessian matrices:

$$J = \begin{bmatrix} \partial x/\partial x \\ \partial x/\partial y \end{bmatrix} = \nabla x$$

$$H = \begin{bmatrix} \partial^2 x/\partial x^2 & \partial^2 x/\partial x \partial y \\ simm. & \partial^2 x/\partial y^2 \end{bmatrix} = \nabla \nabla^t x$$

Therefore given the autocovariance of the stochastic process, the cross-covariance between the first derivatives and the stochastic process itself has the following expression, according to covariance propagation law and the commutability of the operators, because of their linearity:

$$C(\nabla_Q \xi_Q, \xi_P) = \nabla_Q C(r_{PQ}) = C'(r_{PQ}) \nabla_Q (|r_{PQ}|) =$$
$$-C'(r_{PQ}) r_{PQ} / |r_{PQ}| = -C'(\xi_Q, \xi_P) r_{PQ} / |r_{PQ}|$$

being: $r_{QP} = -r_{PQ}$, and: $\nabla|r| = r/|r| = [cosq \quad senq]^t$,where the direction *q* begins from the *x* axis and goes anticlockwise. An important function of the first derivatives is the modulus of the gradient:

$$|\nabla \xi| = (\partial \xi/\partial x)^2 + (\partial \xi/\partial y)^2$$

In the same way, the cross-covariance between the second derivatives of a stochastic process and the stochastic process itself has the following expression:

$$C(\nabla_Q \nabla_Q^t \xi_Q, \xi_P) = \nabla_Q \nabla_Q^t C(r_{PQ}) =$$
$$= -\nabla_Q C'(r_{PQ}) r_{PQ}^t / |r_{PQ}| =$$
$$= C''(r_{PQ}) r_{PQ} r_{PQ}^t / |r_{PQ}|^2 - C'(r_{PQ}) \nabla_Q (r_{PQ} / |r_{PQ}|) =$$
$$= C''(r_{PQ}) \Pi r_{PQ} - C'(r_{PQ}) \left(-I/|r_{PQ}| + r_{PQ} r_{PQ}^t / |r_{PQ}|^3 \right) =$$
$$= (C''(r_{PQ}) - C'(r_{PQ})/|r_{PQ}|) \Pi r_{PQ} + C'(r_{PQ}) I/|r_{PQ}| =$$
$$= (C''(\xi_Q, \xi_P) - C'(\xi_Q, \xi_P)/|r_{PQ}|) \Pi r_{PQ} +$$
$$+ C'(\xi_Q, \xi_P) I/|r_{PQ}|$$

being: $\nabla r^t = I; \quad \nabla(1/|r|) = -\nabla(|r|)/|r^2|^2 = -r/|r|^3$, and: $Pr = rr^t/|r|^2 = \begin{bmatrix} cos^2 q & sinq\ cosq \\ simm. & sin^2 q \end{bmatrix}$, where P represents a projector:
$(Pr)^2 = Pr$; $Tr(Pr) = 1$; $Det(Pr) = 0$. An important function of the second derivatives is the Laplacian:

$$\nabla^2 x = (\partial^2 x/\partial x^2 + \partial^2 x/\partial y^2)$$

that represents an areal dilatation or contraction, whilst the maximum local shear strain is given by:

$$g = \sqrt{(\partial^2 x/\partial x^2 - \partial^2 x/\partial y^2)^2 + (\partial^2 x/(\partial x \partial y))^2/4}$$

Notice that the Taylor-Karman structure belongs also to multidimensional isotropic stochastic process; in particular, an 3D isotropic stochastic process, the vector $\nabla|r|$ and the projector matrix Pr have the following expression:

$$\nabla|r| = [cosq\ cosa \quad sinq\ cosa \quad sina]^t$$

$$Pr = \begin{bmatrix} cos^2 q\ cos^2 a & sinq\ cosq\ cos^2 a & cosqsina\ cosa \\ & sin^2 q\ cos^2 a & sinqsina\ cosa \\ simm. & & sin^2 a \end{bmatrix}$$

being *q* the same angle, as in 2D isotropic stochastic process, and *a* the elevation angle.

Finally, taking into account integral operators to the signal which are also very important, because their linearity and their capability to measure areas, volumes, etc., new functionals can be estimated.
Therefore given the autocovariance of the stochastic process, the cross-covariance between the integral and the stochastic process itself has the following expression, according to covariance propagation law and the commutability of the operators, because of their linearity:

$$C\left(\int_{S_Q} \xi dS_Q, \xi_P \right) = \int_{S_Q} C(\xi, \xi_P) dS_Q$$

Let recall that the domain is an interval, when the function is a line and the integral is an area; on the contrary the domain is an area, when the function is a surface and the integral is a volume. In the multidimensional cases, the domain is a hyperspace, the function a hypersurface and the integral a hypervolume.

Finite element interpolation

Finite element interpolation can be performed in different way; one the most promising method uses spline functions (e.g. cubic spline functions). A spline function, in the space and time domain, is given by the sum of the product of two or three orthogonal cubic spline functions (in the space domain), plus one cubic spline function (in the time domain):

$$S(P,t) = S(x)S(y)S(z) + S(t)$$

The choice for the number of cells and the number of knots depend on the number of observations *m* and the interpolation steps d_p, d_t.
The number of cells is the sum of the product of the number of classes in two or three directions *x* ,*y* and *z*, plus the number of the classes in the time:

$$v = v_x v_y v_z + v_t$$

where: $n_x = int(Dx/d_p) + 1$, $n_y = int(Dy/d_p) + 1$, $n_z = int(Dz/d_p) + 1$, $n_t = int(Dt/d_t) + 1$, being Dx, Dy and Dz the dimensions of the space region in two or three directions and d_p the chosen interpolation step in the space domain, whilst

Dt is the dimension of the time interval and d_t. the chosen interpolation step in the time domain. Consequently the number of knots is:

$$n = n_x n_y n_z + n_t = (v_x + 3)(v_y + 3)(v_z + 3) + (v_t + 3)$$

The spline interpolation is performed, as a classical least squares problem, by writing a system of observation equations:

$$\hat{s}_h = s_h^0 + \hat{v}_h = \sum_{i=1}^{4}\sum_{j=1}^{4}\sum_{k=1}^{4} \hat{a}_{I+i\ J+j\ K+k} S_{ijk}(x_h, h_h, z_h) + \sum_{l=1}^{4} b_{L+l} S_l(t_h)$$

and associating it with the least squares norm:

$$f = \sum_{k=1}^{m} \hat{v}_k^2 = min$$

The weights are mostly assumed equal one; however more complex stochastic model should be defined eventually including correlations among the observations, but they are usually omitted for sake of brevity. The following formulas are the legenda of the functional model. Indeed for the x direction, the coordinate of the h-th knot with respect to the initial corner is split in two parts:

$$Dx_h = Id_p + dx_h$$

where the number of the preceding knots is:

$$I = int(Dx_h / d_p)$$

and the position inside the class is:

$$x_h = dx_h / d_p$$

being: $dx_h = Dx_h - Id_p$. Analogously for the y and z directions:

$$Dy_h = Jd_p + dy_h$$

$$Dz_h = Kd_p + dz_h$$

where: $J = int(Dy_h / d_p)$, $h_h = dy_h / d_p$, $dy_h = Dy_h - Jd_p$ and $K = int(Dz_h / d_p)$, $z_h = dz_h / d_p$, $dz_h = Dz_h - Kd_p$.
Moreover for the time:

$$Dt_h = Ld_t + dt_h$$

where: $L = int(Dt_h / d_t)$, $t_h = dt_h / d_t$, $dt_h = Dt_h - Ld_t$.
Notice that suitable constraints for the knots should be introduced at the border and in empty regions, if any. Indeed the Tikhonov norm induces continuity of the model, by imposition of the regularity of the first derivaties. It solves ill-conditioned subsystems, if any; whilst the low weights used don't destroy the model, in case of well-conditioning.
In practice finite difference equations of the first order are written, for each knot, taking into account: in the space domain, the knots left and right (along the same row), down and up (along the same column), lower and upper (along the same pile), and in the time, the knots preceding and following:

$$a_{I+i+1\ J+\ jK+k} - a_{I+i-1\ J+\ j\ K+k} = 0$$

$$a_{I+i\ J+\ j+1K+k} - a_{I+i\ J+\ j-1\ K+k} = 0$$

$$a_{I+i\ J+\ jK+k+1} - a_{I+i\ J+\ j\ K+k-1} = 0$$

$$b_{L+l+1} - b_{L+l-1} = 0$$

Notice that the pseudo-observations are always equal to zero, being it the median value of the first derivaties. Furthermore the weights are chosen suitably low, so that the constraints solve the problems given by the border and the empty regions, whilst they respect the information supplied by the data.

Acknowledgment

This study has been developed within the project ``Definition of a quality model for 3D computer cartography carried out by digital photogrammetry, with suitable features for representing urban buildings, for planning mobile phone networks". Partly financed by MURST (Italian Ministry of University and Research) in 1997 as a project of relevant national interest. National coordinator: Riccardo Galetto. Head of the research unit: Luigi Mussio.

References

Ammannati, F. & B. Benciolini & Mussio L. & F. Sansò, 1983, An experiment of collocation applied to digital height model analysis. In K. Torlegard (Ed) *Int. Colloquium on Mathematical aspects of digital elevation models*, Dept. of Photogrammetry, Royal Institute of Technology, Stockholm.

Ammannati, F. & B. Betti, & L. Mussio, 1984, Statistical analysis of relevant terrain features through digital elevation models. In *Int. Archives of Photogrammetry and Remote Sensing*, Vol. XXV, Part A3a/b, Rio de Janeiro.

Baldi, P. & B. Benciolini, & Gasperini, P. & L. Mussio, & Sansò, F. & S. Zerbini, 1986, Simulation for the PDN85 doppler campaign. *Bollettino di Geodesia e Scienze Affini*, Vol. 45, n. 1.

Barzaghi, R. & B. Benciolini, & Betti, B. & G. Forlani, & Mussio, L. & F. Sansò, 1990, A numerical experiment of integrated geodesy. *Bulletin Geodesique*, Vol. 64, n. 3.

Barzaghi, R. & B. Crippa, & Forlani, G. & L. Mussio, 1988, Digital modelling by using the integrated geodesy approach. In. *Int. Archives of Photogrammetry and Remote Sensing*, Vol. XXVII, Part B3, Kyoto.

Bellone, T. & B. Crippa, & L. Mussio, 1996, A recommended procedure for image quality assessment. In L. Mussio, G. Forlani (Ed's) *Int. Archives of Photogrammetry and Remote Sensing*, Vol. XXX, Part 1.bis, Milan.

Bellone, T. & B. Crippa, & L. Mussio, 1996, Restoration and Validation of Image Data. In L. Mussio, G. & F. Forlani, Crosilla, (Ed's) *Data Acquisition and Analysis for Multimedia GIS*, CISM Udine, Courses and Lectures - no. 365, Springer, Wien - New York.

Bellone, T. & L. Mussio, & C. Nardinocchi, 1997, On Data Processing in Photogrammetry and Cartography. In Mussio, L. & B. Crippa, & A. Vettore (Ed's) *Int. Archives of Photogrammetry and Remote Sensing*, Vol. XXXII, Part 6W1, Milan.

Benciolini, B. & B. Betti, & Mussio, L. & F. Sansò, 1988, Experiments of Integrated Geodesy in Geodesy and Related Sciences. *Bollettino di Geodesia e Scienze Affini*, Vol. 47, n. 3.

Betti, B. & L. Mussio, & F. Sansò, 1983, A new proposal of sphere reconstitution in Hipparcos project. In P.L. Bernacca (Ed) *The FAST thinkshop Processing of scientific data from the ESA astrometry satellite Hipparcos*, CNR-PSN University of Padua, Padua.

Betti, B. & L. Mussio, & F. Sansò, 1985, Experiments with the arcwise sphere reconstitution. In J. Kovalevsky (Ed) *The 2nd FAST thinkshop Processing of scientific data from the ESA astrometry satellite Hipparcos*, CERGA, Grasse.

Colombo, L. & G. Fangi., & Mussio, L. & F. Radicioni, 1986, Further development on digital models in a control problem for the "Ancona '82" landslide. In E. Kilpelae, A. Savolainen, A. Laiho (Ed's) *Int. Archives of Photogrammetry and Remote Sensing*, Vol. XXVI, Part 3.1, Rovaniemi.

Colombo,L. & G. Fangi, & Mussio, L. & F. Radicioni, 1987, Kinematic processing and spatial analysis of leveling control data of the Ancona '82 landslide. In H. Pelzer and W.Niemeier (Ed's) *Determination of Heights and Height Changes*, Duemmler Verlag, Bonn.

Crespi, M. & B. Crippa, & L. Mussio, 1990, Integrated approach to kinematic adjustment and spatial analysis of leveling control data of the Ancona 82 landslide. In P. Vyskocil, C. Reigber and P.A. Cross (Ed's) *Global and Regional Geodynamics*, Springer-Verlag, New York.

Crippa, B. & M. Crosetto, & L. Mussio, 1998, The Use of Interferometric SAR for Surface Reconstruction. In *Int. Archives of Photogrammetry and Remote Sensing*, Vol. XXXII, Part 1, Bangalore.

Crippa, B. & A. De Haan, & L. Mussio, L.1989, The formal structure of geodetic and photogrammetric observations. In *ISPRS IC-WG III/VI Tutorial on Mathematical Aspects of Data Analysis*, Polytechnic of Milan, Milan.

Crippa, B. & A. De Haan, & L. Mussio, 1992, The creation of a DEM from SPOT data. In *European International Space Year - Space in the service of the changing Earth - Environment Observation and Climate Modelling through Int. Space Projects*, ESA ISY - 1, Paris Cedex.

Crippa, B. & G. Forlani, & L. Mussio, 1993, Digital elevation model of the Torre del Sale (Piombino) Thermal Power Plant area. In Harts, J. & H.F.L. Ottengs, & H.I.J. Scholten (Ed's) *EGIS 93*, Vol. I, EGIS Foundation, Utrecht - Amsterdam.

Crippa, B. & H. G. Maas, & L. Mussio, 1992, Spatial analysis of turbulent flow fields by deterministic and stochastic approaches. In L.W. Fritz, J.R. Lucas (Ed's) *Int. Archives of Photogrammetry and Remote Sensing*, Vol. XXIX, Part B5, Washington D.C.

Crippa, B. & E. S. Malinverni, & L. Mussio, 1995, Dynamic modelling of scene deformation for crustal movement monitoring. In A. Gruen, E. Baltsavias (Ed's) *Int. Archives of Photogrammetry and Remote Sensing*, Vol. XXX, Part 5W1, Zurich.

Crippa, B. & L. Mussio, 1987. The new ITM System of the programs MODEL for digital modelling. In O. Jacoby and P. Frederiksen (Ed's) *Int. Colloquium on Progress in Terrain Modelling*, Technical University of Denmark, Copenhagen.

Crippa, B. & L. Mussio, & Sharif, 1994, The Torre del Sale (Piombino) DEM and its contribution to an Environmental Monitoring oriented GIS. In G. Guariso and B. Page (Ed's) *Computer support for Environmental Impact Assessment*, North Holland, Amsterdam.

Crippa, B. & L. Mussio, & G. Zacco, 1994, Data understanding by form descriptors on the way of compression. In L. Mussio, G. Forlani (Ed's) *Int. Archives of Photogrammetry and Remote Sensing*, Vol. XXX, Part 1.1, Milan.

Cunietti, M. & G. Fangi, & Mussio, L. & F. Radicioni, 1984, Block adjustment and digital model of photogrammetric data in a control problem for the "Ancona '82" landslide. In *Int. Archives of Photogrammetry and Remote Sensing*, Vol XXV, Part A3a/b, Rio de Janeiro.

De Haan, A. & H. G. Maas, & Malik, N. & L. Mussio, 1994,. Statistical analysis of spatially referenced data. *Inst. fuer Geodaesie und Photogrammetrie ETH-Zurich*, Bericht 238.

De Haan, A. & L. Mussio, 1989, Kinematic levelling adjustments with polynomials and polynomial splines. *Geodesia and Scienze Affini*, Vol. 48, n. 3.

Forlani, G. & L. Mussio, 1986, Test on joint adjustment of geodetic and photogrammetric data. In E. Kilpelae, A. Savolainen, A. Laiho (Ed's) *Int. Archives of Photogrammetry and Remote Sensing*, Vol. XXVI, Part 3.1, Rovaniemi.

Forlani, G. & L. Mussio, 1986, The kinematic adjustment of levelling network. In M. Unguendoli (Ed) *Modern trends in deformations measurements* , CUSL, Bologna.

Migliacci, A. & G. Mirabella Roberti, & Monti, C. & L. Mussio, 1986, Monitoring of historical wall painting. In *IABSE Safety and quality assurance of civil engineering structures* , Tokio.

Mussio, L. & F. Radicioni., 1986, High precision photogrammetric measurements to control the shape of a section of the Musone river. In M. Unguendoli (Ed) *Meeting on the Modern trends in deformations measurements* , CUSL, Bologna.