CHALLENGING THE LOGIC OF LEAST-SQUARES METHODS FOR GEODETIC ESTIMATION PROBLEMS

Christopher Kotsakis

Abstract

Least-squares estimation methods are perhaps the most widely used tool in all fields of geodetic research. Nevertheless, their prevailing use is not often complemented by a widespread objective view of their rudiments. Within the standard formalism of the least-squares estimation theory there are actually several paradoxical and curious issues which are seldom explicitly formulated. The aim of this expository paper is to present some of these issues and to discuss their implications for geodetic data analysis and parameter estimation problems.

« De tous les principes qu'on peut proposer pour cet objet, je pense qu'il n'en est pas de plus general, de plus exact, ni d'une application plus facile que celui qui consiste à rendre minimum la somme de carrés des erreurs. »
Adrien-Marie Legendre [1805]

1. Introduction

The method of least-squares (LS) estimation has a long history that spans almost two centuries of human intellectual work. Having been conceived in the nineteenth century when new and revolutionary ideas were systematically emerging in most areas of scientific research, least-squares theory marks the starting point of modern data analysis for the applied sciences. Over the years LS methods have managed to acquire a status of "universal legitimacy" and they still provide the first tool that comes to mind when one deals with optimal prediction or parameter estimation problems. An example attesting to this fact is the formal procedure employed by CODATA (Committee on Data for Science and Technology) to determine the numerical values of the fundamental physical constants. CODATA integrates several measurements which are obtained from different physical experiments and analyzes them with a common least-squares adjustment procedure that yields the best estimates for most fundamental constants. These LS estimates are subsequently suggested for general use by the International Council of Science (formerly ICSU) and they are recognized worldwide for all fields of science and engineering [*Mohr and Taylor* 2001, 2003].

In the geodetic world, the logic of LS estimation has undoubtedly provided a fundamental guiding principle that is routinely applied in every field (photogrammetry, remote sensing, geodetic positioning, geoid determination, deformation monitoring, etc.) where optimal analysis of geo-related or geo-referenced data is needed [*Dermanis et al.* 2000]. Moreover, one can reasonably claim that there is a strong sentimental value hidden in the mutual relationship between least-squares and geodetic data analysis, since it was a geodetic positioning problem which led the German mathematician *Carl Friedrich Gauss* to the formal statistical development of the LS adjustment method.

The prevailing use of LS methods in geodesy, however, has not been complemented by a widespread common understanding of their rudiments. Although there is complete agreement on how to form the so-called *normal equations* from a system of *observation equations* and everyone can obtain the very same results for their solution, the reasons for employing LS-based estimation techniques, the perception of their objectives and the conditions under which these are achieved, as well as the interpretation of their final results, may be quite different among researchers. This somewhat controversial situation is probably due to the fact that least-squares, as a method for optimal data processing and inversion, was originally invented from three distinctively separate perspectives: (i) *least sum of squared residuals* [*Legendre* 1805], (ii) *maximum probability of zero error of estimation* [*Gauss* 1809], and (iii) *least mean squared error of estimation* [*Gauss* 1821, 1823, 1826]; see also *Plackett* [1972].

The previous realizations may seem to be out-of-date and of little concern in today's pragmatic world where often practical results overshadow theoretical quests. Nevertheless, the objective of this paper is to present an alternative viewpoint of the optimal statistical principles that are traditionally linked to LS estimation. The main focus is put on switching the property of *unbiasedness* for the LS estimators with a different, yet equivalent, constraint. In particular, it will be explained how the same optimal LS solution can be obtained if we replace the a-priori requirement of unbiasedness with a condition which implies that the numerical range of the unknown parameters is unbounded. The theoretical and practical consequences that arise from this strange dualism are discussed, and a short critique on the statistical foundations of the least-squares method is also made. To avoid any possible misunderstanding it should perhaps be added that there is nothing mathematically new about the considerations presented in this paper. What we only attempt to convey here is that the logic of the statistical aspects that are typically associated with LS estimation methods receive a more objective and careful treatment than is usually given to them.

2. The usual statistical approach for LS estimation problems

In the context of statistical inference, the LS methodology provides a *linear* and *uniformly unbiased* estimator that has the *minimum mean squared error* among any other linear unbiased estimator. This is the standard perspective that is typically followed to describe the optimality of LS estimation techniques and it has been the basis upon which their choice for the solution of practical problems is usually justified. In a historical context, this probabilistic/frequentist viewpoint is due to *Gauss*'s second formulation for the LS method. The rigorous link between this approach, which will be identified thereafter by the acronym BLUE (Best Linear Unbiased Estimation), and *Legendre*'s deterministic/algebraic formulation (i.e. minimum sum of the squared residuals) is provided by the well known Gauss-Markov theorem [e.g. *Dermanis and Rummel* 2000, pp. 48-49].

In this section a short review of the BLUE version for the LS estimation process is given. Although our presentation does not follow the most general mathematical setting, it is nevertheless sufficient for the purpose of this paper.

2.1 The general linear model

Let us begin with a system of linear(-ized) observation equations which has the general form given in Eq. (1). Note that this is an exceptionally convenient and compact setting since most of the geomatics applications that are related to optimal data analysis (e.g. geodetic network adjustment, image analysis, satellite orbit modelling, gravity field determination, deformation monitoring, etc.)

can essentially be reduced to an inversion problem for such a system of equations. In general terms, we thus have

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{v} \tag{1}$$

where **y** is a known observation vector, **x** is an unknown parameter vector and **A** is a design matrix of known coefficients with full column rank. The residual vector **v** contains unknown random errors (data noise) whose statistical characteristics are typically given in terms of their first and second order moments, i.e.

$$E\{\mathbf{v}\} = \mathbf{0}\ , \quad E\{\mathbf{v}\mathbf{v}^{\mathrm{T}}\} = \mathbf{C} \tag{2}$$

The symbol $E\{\cdot\}$ denotes the mathematical expectation operator (i.e. average taken over all possible repetitions of the measurement set). In practice, the error covariance (CV) matrix **C** is often considered partially known and its uncertainty is commonly controlled by one or more scaling factors (variance components) which can be estimated a-posteriori from the available data. Since the knowledge of the error CV matrix does not play a crucial role in the rest of this paper, we assume that **C** is a fully known symmetric and positive-definite matrix.

The linear model of Eq. (1), along with the stochastic error description of Eq. (2), is suitable for the study of a variety of physical systems, including most of the fields in modern geodetic research. In principle, in all such cases we generally seek to estimate an unknown quantity θ which depends on the parameter vector **x**. For convenience, we consider here the case where θ is a linear function of the unknown parameters

$$\theta = \mathbf{q}^{\mathrm{T}}\mathbf{x} \tag{3}$$

with **q** being an arbitrary known vector. Based on the knowledge of the data vector **y**, various types of estimators $\hat{\theta} = \hat{\theta}(\mathbf{y})$ can be considered, each of which complies with specific optimal criteria and assumptions. In the following, we describe the statistical characteristics of the classic LS estimators that can be associated with the aforementioned linear model.

2.2 Least-squares solution as a BLUE estimator

A general *linear estimator* of the unknown quantity θ will have the form

$$\hat{\theta} = \mathbf{b}^{\mathrm{T}}\mathbf{y} + \mathbf{c} \tag{4}$$

where the vector **b** and the scalar *c* need to be determined according to some optimality criteria. The statistical formulation of LS estimation is based on two fundamental properties that the linear estimator of Eq. (4) should satisfy simultaneously, namely

- *Uniform unbiasedness* – $E\{\hat{\theta}\} = \theta = \mathbf{q}^{\mathrm{T}}\mathbf{x}$ for any parameter vector **x**

- *Minimum mean squared error* – $E\{(\hat{\theta} - \theta)^2\} = \text{minimum}$

It can easily be shown that the first property leads to the following constraints for the quantities **b** and c

$$\mathbf{b}^{\mathrm{T}}\mathbf{A} = \mathbf{q}^{\mathrm{T}} \tag{5}$$

and

$$c = 0 \tag{6}$$

Also, using Eqs. (3) and (4) we can establish that the mean squared error (MSE) of the linear estimator $\hat{\theta}$ has the general form

$$E\{(\hat{\theta}-\theta)^2\} = \mathbf{b}^{\mathrm{T}}\mathbf{C}\mathbf{b} + [(\mathbf{b}^{\mathrm{T}}\mathbf{A}-\mathbf{q}^{\mathrm{T}})\mathbf{x}+c]^2 \tag{7}$$

The minimization of the above quantity, in conjunction with the linear constraints of Eqs. (5) and (6), lead to a unique optimal solution for **b** through the method of Lagrange multipliers. The result is given by the following equation

$$\mathbf{b} = \mathbf{C}^{-1}\mathbf{A}(\mathbf{A}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{A})^{-1}\mathbf{q} \tag{8}$$

Based on Eqs. (4), (6) and (8), the LS estimate of $\theta = \mathbf{q}^{\mathrm{T}}\mathbf{x}$ is thus given by the well known expression

$$\hat{\theta} = \mathbf{q}^{\mathrm{T}}(\mathbf{A}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{y} \tag{9}$$

which, in turn, implies the following LS estimate for the parameter vector **x**

$$\hat{\mathbf{x}} = (\mathbf{A}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{y} \tag{10}$$

In many textbooks the statistical optimality of the LS method is normally attributed to the fact that it provides the minimum *error variance* among all other linear unbiased estimation algorithms. Due to the unbiasedness property of the LS estimator $\hat{\theta}$, however, the variance and the mean squared value of its estimation error $\hat{\theta}-\theta$ are exactly equal. As a result, the BLUE formulation can be equivalently based either on the minimization of the *error variance* or on the minimization of the *mean squared error* for a linear unbiased estimator. Here we have chosen to follow the latter approach since it provides a more direct connection with the discussion given in the following section.

3. An alternative view of the LS estimation process

The same estimators given in Eqs. (9) and (10) can be also obtained through a different formulation, without however departing from the broad context of optimal statistical inference. The alternative approach that is presented here represents only an attempt to elucidate the logic of the unbiasedness condition which is associated with LS estimators.

If we keep the same setting that was adopted in the last section and start from a typical linear estimator $\hat{\theta} = \mathbf{b}^{\mathrm{T}}\mathbf{y} + c$ for an unknown scalar quantity $\theta = \mathbf{q}^{\mathrm{T}}\mathbf{x}$, we again seek optimal values for **b** and *c*. As it was mentioned already, the mean squared estimation error in such a case has the general form

$$E\{(\hat{\theta} - \theta)^2\} \;=\; \mathbf{b}^{\mathrm{T}}\mathbf{C}\mathbf{b} \;+\; [(\mathbf{b}^{\mathrm{T}}\mathbf{A} - \mathbf{q}^{\mathrm{T}})\mathbf{x} + c]^2$$

where **A** is the design matrix of the linear(-ized) system of observation equations; see Eq. (1).

Let us now indicate the critical fact that the MSE of $\hat{\theta}$ depends directly on the vector of the original unknown parameters, as it is clearly seen from the last formula. Consequently, *if the range of* **x** *is unbounded*, the second term in the above expression becomes, in general, unbounded too. In order to ensure that the MSE of the linear estimate $\hat{\theta}$ remains finite, regardless of the numerical range of the unknown parameters, the following condition should thus be satisfied

$$\mathbf{b}^{\mathrm{T}}\mathbf{A} \;-\; \mathbf{q}^{\mathrm{T}} \;=\; \mathbf{0}^{\mathrm{T}} \qquad (11)$$

where $\mathbf{0}^{\mathrm{T}}$ corresponds to a row vector of zeros. Subject to this condition and given the fact that *c* is only a constant scalar, the MSE minimization for the linear estimator $\hat{\theta}$ yields the result

$$\mathbf{b} \;=\; \mathbf{C}^{-1}\mathbf{A}(\mathbf{A}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{A})^{-1}\mathbf{q} \quad \text{and} \quad c \;=\; 0 \qquad (12)$$

which, in turn, gives rise to the same estimates for θ and **x** that were derived in the previous section.

Hence, it is seen that we can obtain the same optimal solution as in the BLUE case, without invoking a-priori the requirement of having unbiased results for the estimated parameters. An equivalent formulation of the LS estimation process can thus emerge which is summarized as follows: *among all linear estimators that provide finite mean squared error for a set of unknown parameters with unbounded range, least-squares estimators yield results with minimum mean squared error*.

Under the preceding perspective it may appear that we have removed the requirement of unbiasedness at the expense of a more restricted version for the LS method. Obviously, the property of unbiasedness for the optimal solution has not been lost in this case, since it will now be a direct consequence of Eq. (12). On the other hand, the resulting estimators are not restrictive in any way because they can be always implemented, regardless of the actual range of the parameter vector **x** and/or the numerical values in the data vector **y**. In fact, what the previous alternative formulation should make us skeptical about is the following question: *will the traditional LS estimation algorithms give optimal results in the case where* **x** *is a vector of bounded parameters ?*

4. Discussion - Conclusions

An informative way to look at LS estimation is to recognize that its statistical optimality is closely associated with the assumption that the range of the unknown parameter vector **x** is unbounded. Clearly, in all geodetic applications where LS techniques are employed, the values of the parameters that need to be estimated always lie within a finite range. What should be acknowledged here is that this important piece of information (or perhaps a *fact* for most physical systems under study) is not integrated at all in the ordinary LS estimation process. The statistical logic of the LS principle, as this is depicted in terms of a linear unbiased estimator with minimum MSE, ignores the fact that the unknown parameters are always finite in magnitude. That is probably the reason why LS solutions tend to give numerical answers that are "longer" (when measured by some Euclidean-type norm) than the actual true parameter vector. It is actually instructive to recall the following well known formula from LS adjustment theory [*Sen and Srivastava* 1990]

$$E\left\{\hat{\mathbf{x}}^{\mathrm{T}}\hat{\mathbf{x}}\right\} = \mathbf{x}^{\mathrm{T}}\mathbf{x} + \mathrm{trace}(\mathbf{A}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{A})^{-1} \tag{13}$$

which indeed shows that the LS-estimated parameter vector $\hat{\mathbf{x}}$ is expected to be always longer than the corresponding true vector $\mathbf{x}$ (i.e. the matrix $\mathbf{A}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{A}$ is positive-definite and thus the trace of its inverse is always a positive number).

An even more interesting and important conclusion which can be drawn from the comparative analysis of sections 2 and 3 is that the property of unbiasedness is responsible for causing the LS estimators to be "blind" on the bounded nature of the unknown parameters. This rather strange dualism brings up a fairly strong argument in favor of biased estimation methods, although in geodesy we have often been guided by a non-enthusiastic, if not negative, attitude towards the use of such techniques; for a general overview of biased estimation methods see *Mayer and Willke* [1973] and for a comparison between biased and unbiased estimation techniques see *Efron* [1975]. But, on the other hand, is it reasonable to embrace a statistical estimation philosophy whose one of its optimality principles is associated with the assumption that the unknown parameters are boundless? Should we not prefer an estimation algorithm which respects the fact that physical quantities, such as the Cartesian coordinates in geodetic/surveying networks, the transformation parameters between spatial reference systems or the deformation rates in geotechnical structures (to name a few of the usual unknowns that appear in geodetic applications), cannot exceed certain limits? Is it not reasonable to adjust a levelling network by incorporating the prior information that the unknown height values of its points do not exceed, say, 6,000 m? We should also not forget that the linear model in Eq. (1), which is extensively used in most applications of geodetic data analysis, gives a realistic picture of physical reality only when the unknowns $\mathbf{x}$ are constrained within certain intervals around their initial approximate values [*Björck* 1996, p. 195].

At this point, one can rightly protest that the finite range of the unknown parameters can easily be taken into account in practice, without giving up the standard LS principle. That is, we can always seek optimal estimators $\hat{\mathbf{x}}$ which minimize the sum of the squared residuals in the misclosure vector $\mathbf{y} - \mathbf{Ax}$ (*Legendre's* formulation for the LS method), subject to appropriate *inequality constraints* that bound the size of the unknown parameters $\mathbf{x}$. Such problems are actually very common in various disciplines of engineering and geosciences (tomographic geophysical inversions, ocean circulation modelling, optimization of mechanical systems, etc.) and they can be handled through well known techniques of convex optimization and non-linear programming [e.g. *Luenberger* 1969, 1984]. What is important however to note here is that the solutions of such constrained LS adjustments, using either linear or quadratic inequality constraints, will not generally produce unbiased estimates for the unknown parameters.

As an example, we can mention a relatively common problem of constrained optimization, according to which an upper bound $b_{\max}$ is placed on the Euclidean length of the parameter vector $\mathbf{x}$ that appears in the observation equations model of Eq. (1). If we now seek the least-squares estimator (in the Legendre's sense - $(\mathbf{y}-\mathbf{Ax})^{\mathrm{T}}\mathbf{C}^{-1}(\mathbf{y}-\mathbf{Ax}) = minimum$) subject to the inequality constraint $\mathbf{x}^{\mathrm{T}}\mathbf{x} \le b_{\max}$, the final result will have the *ridge regression* form

$$\hat{\mathbf{x}} = (\mathbf{A}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{A} + k\mathbf{I})^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{y} \tag{14}$$

where the scalar parameter k is uniquely determined from the bound $b_{\max}$ of the parameter vector length; for more details and proofs see *Meeter* [1966], *Draper and Smith* [1998, pp. 391-395], *Björck* [1996, p. 205-206]. Taking into account the stochastic model for the data errors in Eq. (2), it is easily verified that the previous linear estimator is generally biased, i.e.

$$E\{\hat{\mathbf{x}}\} = (\mathbf{A}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{A} + k\mathbf{I})^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{Ax} \neq \mathbf{x} \tag{15}$$

In fact, the solution in Eq. (14) corresponds to a well known estimation algorithm which is often viewed as the outcome of a *Tikhonov regularization* procedure for the optimal solution of inverse problems; see e.g. *Bouman and Koop* [1997]. The above estimator is actually known to outperform, in the MSE sense, the ordinary (unconstrained and unbiased) LS solution from Eq. (10) when $\mathbf{x}^{\mathrm{T}}\mathbf{x}$ is bounded [*Marquardt* 1970, *Hoerl and Kennard* 1970]. When viewed under this perspective, it is difficult to argue against the use of ridge regression or other biased estimation methods that confine reasonably the size of the unknown parameters (e.g. "shrunken" estimators), although in reality **x** and its exact length are always unknown.

We should also bring to the reader's attention the fact that the unbiasedness criterion is intrinsically related to the *frequentist* approach in probability theory. The latter provides the backbone of what is known as the "objective" or "classic" view for statistical inference problems, in contrast to the more controversial *Bayesian* methods which represent the movement of "subjectivists" in estimation theory [*Jaynes* 2003]. Note that for true Bayesians there is usually no contemplation about possible biases in their estimators since this notion is totally irrelevant within the Bayesian vision of statistical inference [*Barnett* 1982, pp. 16-19, 225-226]. An interesting argument against the statistical logic of least-squares methods can be now raised, since the *asymptotic* behaviour of an estimator is not really relevant for practical purposes. Let us not forget that the merit of the unbiasedness principle is the errorless recovery of the true unknown parameters after infinitely many repetitions of the experimental process. However, the real problem in all applied scientific disciplines has always been to achieve the best inference we can from a <u>particular</u> and finite set of data values. Adopting a frequentist's approach who fantasizes about continuously repeated measurements and optimizes the average estimation performance over infinitely many data sets that are obtained under identical conditions can be considered irrelevant, since it corresponds to an imaginary scenario that is never attained in practice.

As a final remark, we should say that this paper does not represent a disagreement with the practice of applying the LS methodology for optimal estimation problems. What is merely claimed here is that the statistical logic of least-squares estimation receives a more objective treatment than it is usually given. In particular, the widespread belief that unbiasedness is a natural, almost purifying property that should accompany any meaningful optimal estimator must be rationally criticized, in view of its strong link with the assumption that the unknown parameters are boundless. In this way, for example, potential users of ridge regression or other similar estimation techniques that compete with the standard BLUE-type methods would not think (only) in terms of having to choose between a biased or an unbiased estimator, but they would also consider the concurrent knowledge about physical reality that each of these methods brings to an investigator's quest for understanding.

References

Barnett, V. 1982. *Comparative statistical inference*. John Wiley and Sons Ltd, 2nd edition.

Björck, A. 1996. *Numerical Methods for Least Squares Problems*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia.

Bouman, J. and R. Koop. 1997. Quality differences between Tikhonov regularization and generalized biased estimation. *DEOS Progress Letter*, 97, 1, pp. 42-48.

Dermanis, A. and R. Rummel. 2000. Data analysis methods in geodesy. In: *A. Dermanis et al.* [2000], pp. 17-92.

Dermanis, A., A. Gruen and F. Sanso (editors). 2000. *Geomatic Methods for the Analysis of Data in Earth Sciences*. Lecture Notes in Earth Sciences Series, vol. 95, Springer Verlag, Berlin Heidelberg.

Draper, N.R. and H. Smith. 1981. *Applied Regression Analysis*. 2nd edition, Wiley, New York.

Efron, B. 1975. Biased versus unbiased estimation. *Advances in Mathematics*, 16, pp. 259-277.

Gauss, C.F. 1809. *Theoria Motus Corporum Coelestium in Sectionibus Conicis Solem Ambientium*. Perthes and Besser, Hamburg. English translation (originally in 1857 by C.H. Davis) reprinted as *Theory of the Motions of the Heavenly Bodies Moving about the Sun in Conic Sections*, Dover, New York, 1963.

Gauss, C.F. 1821. Theoria combinationis observationum erroribus minimis obnoxiae: Pars prior. *Göttingische gelehrte Anzeigen*, 33, pp. 321-327.

Gauss, C.F. 1823. Theoria combinationis observationum erroribus minimis obnoxiae: Pars posterior. *Göttingische gelehrte Anzeigen*, 52, pp. 313-318.

Gauss, C.F. 1826. Supplementum theoriae combinationis observationum erroribus minimis obnoxiae. *Göttingische gelehrte Anzeigen*, 153, pp. 1521-1527.

Hoerl, A.E. and R.W. Kennard. 1970. Ridge Regression: Biased Estimation for Nonorthogonal Problems. *Technometrics*, 12, 1, pp. 55-67.

Jaynes, E.T. 2003. *Probability theory – The logic of science*. Cambridge University Press, Cambridge.

Legendre, A.M. 1805. *Nouvelles Méthodes pour la Détermination des Orbites des Comètes* (appendix: *sur la méthode des moindres quarrés*). Courcier, Paris. Cited in *Stigler* [1986] where the appendix on least squares is reproduced.

Luenberger, D.G. 1969. *Optimization by vector space methods*. John Wiley and Sons, New York.

Luenberger, D.G. 1984. *Linear and nonlinear programming*. 2nd edition, Addison Wesley, Reading Massachusetts.

Marquardt, D.W. 1970. Generalized inverses, ridge regression, biased linear estimation and non-linear estimation. *Technometrics*, 12, 3, pp. 591-612.

Mayer, L.S. and T.A. Willke. 1973. On biased estimation in linear models. *Technometrics*, 15, 3, pp. 497-508.

Meeter, D.A. 1966. On a theorem used in nonlinear least-squares. *SIAM Journal on Applied Mathematics*, 14, 5, pp. 1176-1179.

Mohr, P.J. and B.N. Taylor. 2001. Adjusting the values of the fundamental constants. *Physics Today*, 54, 3, pp. 29-34.

Mohr, P.J. and B.N. Taylor. 2003. The fundamental physical constants. *Physics Today*, 56, 8, pp. 6-13.

Plackett, R.L. 1972. The discovery of the method of least squares. *Biometrika*, 59, 2, pp. 239-251.

Sen, A. and M. Srivastava. 1990. *Regression Analysis: theory, methods and applications*. Springer-Verlag, New York.

Stigler, S.M. 1986. *The History of Statistics*. Belknap Press, Cambridge, Massachusetts.

Affiliation:

Dr. Christopher Kotsakis
Assistant Professor
Department of Geodesy and Surveying
Aristotle University of Thessaloniki
University Box 440
Thessaloniki, GR – 54124, Greece

Phone: (+30) 2310-996150
Fax: (+30) 2310-995948
Email: kotsaki@topo.auth.gr