Mathematical Geology, Vol. 39, No. 2, February 2007 (© 2007)
DOI: 10.1007/s11004-006-9070-8

Hybrid Estimation of Semivariogram Parameters[1]

Hao Zhang[2] and Dale L. Zimmerman[3]

Two widely used methods of semivariogram estimation are weighted least squares estimation and maximum likelihood estimation. The former have certain computational advantages, whereas the latter are more statistically efficient. We introduce and study a "hybrid" semivariogram estimation procedure that combines weighted least squares estimation of the range parameter with maximum likelihood estimation of the sill (and nugget) assuming known range, in such a way that the sill-to-range ratio in an exponential semivariogram is estimated consistently under an infill asymptotic regime. We show empirically that such a procedure is nearly as efficient computationally, and more efficient statistically for some parameters, than weighted least squares estimation of all of the semivariogram's parameters. Furthermore, we demonstrate that standard plug-in (or empirical) spatial predictors and prediction error variances, obtained by replacing the unknown semivariogram parameters with estimates in expressions for the ordinary kriging predictor and kriging variance, respectively, perform better when hybrid estimates are plugged in than when weighted least squares estimates are plugged in. In view of these results and the simplicity of computing the hybrid estimates from weighted least squares estimates, we suggest that software that currently estimates the semivariogram by weighted least squares methods be amended to include hybrid estimation as an option.

KEY WORDS: consistency, geostatistics, kriging, maximum likelihood estimation, weighted least squares estimation.

INTRODUCTION

Parametric estimation of the semivariogram or covariogram is a classical problem in geostatistics. One widely used approach to this problem involves fitting the chosen parametric model to the sample (empirical) semivariogram (or covariogram) by a least squares method. Specific least squares methods that have been proposed include ordinary least squares (Journel and Huijbregts, 1978; Clark, 1979), weighted least squares (Cressie, 1985), and generalized least squares (Genton, 1998). Another popular approach is likelihood-based; that is, estimates

[1]Received 7 October 2005; accepted 10 August 2006; Published online: 14 April 2007.
[2]Department of Statistics, Washington State University, Pullman, WA 99164-3144, USA; e-mail: zhanghao@wsu.edu.
[3]Department of Statistics and Actuarial Science, University of Iowa, Iowa City, IA 52242, USA; e-mail: dzimmer@stat.uiowa.edu.

of semivariogram or covariogram parameters are obtained by maximizing the joint likelihood function (or some variation thereof) of the observed data; for examples see Kitanidis (1983), Mardia and Marshall (1984), Zimmerman (1989), and Curriero and Lele (1999). Still another approach is Bayesian, in which a prior distribution for the semivariogram or covariogram parameters is updated by conditioning on the observed data (Handcock and Stein, 1993; Ecker and Gelfand, 1997, 1999). The advantages of least squares methods relative to likelihood-based and Bayesian methods are (a) their computational simplicity, and (b) their availability within widely used geostatistical software packages. Their disadvantages are their less firm theoretical statistics foundation and their inefficiency, as empirical studies (Zimmerman and Zimmerman, 1991; Dubin, 1994) suggest that least squares estimators do not perform as well (in terms of mean squared errors, for example) as likelihood-based estimators.

Often in practice, the estimation of semivariogram parameters is not an end in itself, but is merely the first step towards the more important goal of spatial interpolation (prediction). For the purpose of optimal interpolation, it turns out that some semivariogram parameters are more important to estimate well than others (Stein and Handcock, 1989; Zhang, 2004). For example, in the case of a Gaussian random field with isotropic exponential semivariogram $\gamma(h) = \sigma^2[1 - \exp(-h/\alpha)]$, the ratio σ^2/α affects interpolation more than either the sill parameter σ^2 or range parameter α do individually. In fact, under an infill (i.e. fixed-domain) asymptotic framework, only this ratio affects the interpolation, in the sense that an incorrect semivariogram that has a correct sill-to-range ratio will produce asymptotically the same interpolations as those obtained using the correct semivariogram. Therefore, at least for the exponential model, it is important to estimate this ratio precisely, and it is not unreasonable to expect that this ratio may be of similar importance for some other semivariogram models.

The purpose of this article is to introduce and study a "hybrid" semivariogram estimation procedure that combines weighted least squares estimation of the range parameter with maximum likelihood estimation of the sill (and nugget), in such a way that the sill-to-range ratio in an exponential semivariogram is estimated consistently under infill asymptotics. We will show empirically, for several different semivariogram models, that such a procedure is nearly as efficient computationally, and more efficient statistically for some parameters, than weighted least squares estimation of all of the semivariogram's parameters (though it is still discernibly less efficient statistically than maximum likelihood estimation of all the parameters). Furthermore, we will demonstrate that standard plug-in (or empirical) spatial predictors, which are obtained by replacing the unknown semivariogram parameters with estimates in expressions for the ordinary kriging predictor and kriging variance, are better when hybrid estimates, rather than weighted least squares estimates, are plugged in. Our aim is to convince practitioners who currently use weighted least squares estimates (possibly because they lack access to easy-to-use software for maximum likelihood estimation of semivariogram parameters) that hybrid estimation yields worthwhile benefits in estimation and prediction performance, yet is computationally very fast and easy.

The remainder of the article is organized as follows. First, we develop the idea of hybrid estimation. Next, we present results of an empirical study comparing the performance of the hybrid and weighted least squares estimators. Finally, some concluding remarks complete the article.

HYBRID SEMIVARIOGRAM ESTIMATION

Consider a stationary and isotropic Gaussian random field, $\{Z(\mathbf{s}) : \mathbf{s} \in D\}$, where D is a bounded region in $\mathcal{R}^2$. Let μ denote the mean, and $\gamma(h;\boldsymbol{\theta})$ the semivariogram, of the random field. Here, h is Euclidean distance between sites and $\boldsymbol{\theta}$ is a vector of unknown parameters. We assume initially that the semivariogram is of the form

$$\gamma(h;\boldsymbol{\theta}) = \sigma^2[1 - \rho(h;\alpha)] \tag{1}$$

where $\boldsymbol{\theta} = (\sigma^2, \alpha)'$, σ^2 is a sill parameter, $\rho(\cdot)$ is a continuous correlation function, and α is a range parameter. We suppose that one realization of the process is observed at n sites $\{\mathbf{s}_i \in D : i = 1, \ldots, n\}$ yielding observations $(Z(\mathbf{s}_1), \ldots, Z(\mathbf{s}_n))' = \mathbf{Z}$, and that the inferential objectives are to use these observations to estimate σ^2 and α and predict (interpolate) Z at sites where it was not observed. The interpolation method we consider here is standard empirical best linear unbiased prediction, in which estimates are "plugged in" for semivariogram parameters in the expression for the ordinary kriging predictor. It remains to consider how to estimate the semivariogram parameters.

We motivate the notion of hybrid estimation of these parameters by considering the behavior of parameter estimates and interpolations in the context of an infill sampling scheme, in which the spatial domain is fixed and the maximum distance between any two sampled sites decreases to zero as n gets arbitrarily large. Within this asymptotic framework, Stein (1988, 1990) showed that two different semivariograms could, under certain conditions, yield asymptotically equivalent ordinary kriging interpolations. One important and somewhat surprising implication of Stein's results for model (1) is that it is not necessary to estimate σ^2 and α well individually in order to obtain good interpolations; in fact, in the case of an exponential semivariogram $\gamma(h;\sigma^2,\alpha) = \sigma^2[1 - \exp(-h/\alpha)]$ for example, all that is required for asymptotically optimal interpolation is consistent estimation of the sill-to-range ratio $\tau \equiv \sigma^2/\alpha$. Moreover, Zhang (2004) showed, for the same model, that consistent estimators of σ^2 and α do not exist, but the ratio $\tau = \sigma^2/\alpha$ can be consistently estimated. (We remind the reader that the context for these consistency results is infill asymptotics.) In particular, if the likelihood function for this model is denoted by $L(\sigma^2, \alpha; \mathbf{Z})$ and if we define, for any given value α_0 of α,

$$\hat{\sigma}^2_{ML|\alpha_0} = \text{argmax} \ \ L(\sigma^2, \alpha_0; \mathbf{Z})$$

then $\hat{\sigma}^2_{ML|\alpha_0}/\alpha_0$ is a consistent estimator of τ (Zhang, 2004, Theorem 3). It is easily shown that $\hat{\sigma}^2_{ML|\alpha_0}$ can be expressed in closed form, i.e.,

$$\hat{\sigma}^2_{ML|\alpha_0} = (1/n)\mathbf{Z}'\mathbf{R}^{-1}_{\alpha_0}\mathbf{Z} \tag{2}$$

where $\mathbf{R}_{\alpha_0}$ is the (correlation) matrix with (i, j)th element given by $\rho(\mathbf{s}_i - \mathbf{s}_j; \alpha_0)$.

In fact, Zhang's (2004) infill consistency results hold more generally for random fields with nonzero mean and any correlation function belonging to the Matérn family,

$$\rho(h; \alpha) = \frac{(h/\alpha)^\nu}{\Gamma(\nu)2^{\nu-1}} K_\nu(h/\alpha), \quad h \geq 0$$

where ν is a known smoothness parameter and K_ν is the modified Bessel function of order ν. For this family, a consistent estimator of $\sigma^2/\alpha^{2\nu}$ is given by Zhang (2004), who showed that this ratio is more important to interpolation than each individual parameter. However, the extent to which the results generalize to models with a correlation function other than Matérn is not yet known. Nevertheless, we propose the general estimation of σ^2, for a given α_0, by a hybrid approach. Specifically, we propose to estimate σ^2 by an estimator analogous to (2), i.e.

$$\hat{\sigma}^2_{ML|\alpha_0} = (1/n)(\mathbf{Z} - \hat{\mu}\mathbf{1})'\mathbf{R}^{-1}_{\alpha_0}(\mathbf{Z} - \hat{\mu}\mathbf{1})$$

where $\hat{\mu} = (\mathbf{1}'\mathbf{R}^{-1}_{\alpha_0}\mathbf{1})^{-1}\mathbf{1}'\mathbf{R}^{-1}_{\alpha_0}\mathbf{Z}$ and $\mathbf{R}_{\alpha_0}$ is defined as it was above.

It remains to consider how to choose α_0. Although this choice of α_0 does not affect the consistency of the aforementioned estimator of τ when the correlation model is Matérn, and likewise may not affect the estimator's asymptotic efficiency [as shown by Ying (1991) for the one-dimensional exponential case], it may affect the efficiency of the estimator for a finite sample. Therefore, we do not recommend choosing α_0 arbitrarily. Instead, we propose to use the weighted least squares estimate of α, denoted by $\hat{\alpha}_{WLS}$. The hybrid estimator of $\boldsymbol{\theta}$ is then defined as $\hat{\boldsymbol{\theta}}_H = (\hat{\sigma}^2_H, \hat{\alpha}_{WLS})'$ where $\hat{\sigma}^2_H = \hat{\sigma}^2_{ML|\hat{\alpha}_{WLS}}$, and the hybrid estimator of τ is defined as $\hat{\tau}_H = \hat{\sigma}^2_H/\hat{\alpha}_{WLS}$. Note that due to the closed-form representation of $\hat{\sigma}^2_H$, hybrid estimation requires negligible additional effort or computing time beyond that required to obtain the weighted least squares estimates.

Hybrid estimation can be extended to models with a nugget effect as follows. Suppose the semivariogram model is given by

$$\gamma(h;\boldsymbol{\theta}) = \delta I_{\{h>0\}} + \sigma^2(1 - \rho(h;\alpha)),$$

where $I_{\{h>0\}}$ denotes the indicator function which is equal to 1 if $h > 0$, and 0 otherwise; $\rho(h;\alpha)$ is a nuggetless correlation function defined as in (1); and δ and σ^2 are the nugget effect and partial sill parameters, respectively. The hybrid estimators for δ, σ^2 and the mean μ are defined as the maximum likelihood estimators when α is fixed at its weighted least squares estimate. The maximum likelihood estimates can be obtained iteratively by Fisher scoring (Mardia and Marshall, 1984). Although hybrid estimation in this context requires more computation than it does for a nuggetless model, the computational burden is still much less than that required by maximum likelihood estimation of all the parameters in this context.

PERFORMANCE EVALUATION

In order to compare the performance of hybrid semivariogram estimators to weighted least squares, we carried out a simulation study involving Gaussian processes with four distinct semivariogram models. The four models were as follows:

- Exponential model without nugget

$$\gamma_E(h;\boldsymbol{\theta}) = \sigma^2[1 - \exp(-h/\alpha_E)], \quad h \geq 0$$

- Spherical model without nugget

$$\gamma_S(h;\boldsymbol{\theta}) = \begin{cases} \sigma^2[1.5\left(\frac{h}{\alpha_S}\right) - 0.5\left(\frac{h}{\alpha_S}\right)^3], & \text{for } 0 \leq h \leq \alpha_S \\ \sigma^2, & \text{for } h > \alpha_S \end{cases}$$

- Rational quadratic model without nugget

$$\gamma_R(h;\boldsymbol{\theta}) = \frac{19\sigma^2 h^2/\alpha_R}{1 + 19h^2/\alpha_R}, \quad h \geq 0$$

- Exponential model with nugget

$$\gamma_{EN}(h;\boldsymbol{\theta}) = \delta I_{\{h>0\}} + \sigma^2[1 - \exp(-h/\alpha_E)], \quad h \geq 0$$

We note that the square root of α_R in the rational quadratic model is the effective range. In the first three models, four values of range parameter ($\alpha_E = 0.1, 0.2, 0.3, 0.4$; $\alpha_S = 0.2, 0.4, 0.6, 0.8$; and $\alpha_R = 0.04, 0.16, 0.36, 0.64$) were chosen so as to yield processes with relatively weak to relatively strong spatial correlation, and in γ_{EN} three values of nugget ($\delta = 0.5$, 1.0, 2.0) were considered so as to result in processes with a representative span of nugget-to-sill ratios. Without loss of generality, σ^2 was taken to equal 2.0 in all three models. Finally, α_E in $\gamma_{EN}(\cdot)$ was set equal to 0.2.

For the first three models, we considered three sample sizes ($n = 100, 225, 400$), while for the more computationally demanding fourth model we considered only $n = 100$. For each combination of model, sample size, and parameter value, we obtained 1000 simulated realizations from a zero-mean Gaussian process on the unit square at sites forming a square grid. In the case $n = 100$ the grid locations were $(0.05, 0.15, \ldots, 0.95) \times (0.05, 0.15, \ldots, 0.95)$; in the case $n = 225$ the grid locations were $(1/16, 2/16, \ldots, 15/16) \times (1/16, 2/16, \cdots, 15/16)$; and in the case $n = 400$ the grid locations were $(0.025, 0.05, \ldots, 0.975) \times (0.025, 0.05, \ldots, 0.975)$. Note that the spacing among grid locations as the sample size increases conforms to an infill, rather than an increasing domain, asymptotic framework.

For each simulated dataset, the mean and semivariogram parameters were estimated by the weighted least squares and maximum likelihood methods, and hybrid estimators of the sill σ^2, ratio $\tau = \sigma^2/\alpha$, and nugget δ (for the third model only) were obtained. Estimation performance was measured by the empirical bias and mean squared error (MSE) of the estimators over the 1000 simulations. Results are displayed in Tables 1–4. Because our interest lies mainly in comparing the performance of hybrid estimation to that of weighted least squares, we give results only for those parameters whose hybrid and weighted least squares estimators are distinct. Furthermore, we give results on maximum likelihood estimation for the first model only, as it is already known (from theoretical considerations) that maximum likelihood estimation is superior to the other two methods and we merely want to give some indication of how much better it is.

For the exponential model without nugget (Table 1), the hybrid estimator of the sill performed slightly better than its weighted least squares counterpart with respect to bias, but not with respect to MSE. However, the hybrid estimator of the ratio τ was markedly superior to its weighted least squares counterpart with respect to both bias and MSE. As expected, the maximum likelihood estimators of sill and ratio performed much better than the weighted least squares and hybrid estimators across all sample sizes and values of the range parameter. For the spherical model without nugget (Table 2), the hybrid estimators of the sill and ratio had about the same bias as their weighted least squares counterparts, but smaller mean squared errors. For the rational quadratic model without nugget (Table 3), the hybrid estimator of the sill was inferior to the

Table 1. Empirical Bias and MSE of Semivariogram Parameter Estimates for Exponential Model without Nugget

		Bias for sill			MSE for sill		
n	α_E	WLS	Hyb	ML	WLS	Hyb	ML
100	0.1	0.04 (0.01)	0.03 (0.01)	−0.01 (0.01)	0.21	0.23	0.15
	0.2	0.27 (0.04)	0.25 (0.04)	0.01 (0.02)	1.84	1.84	0.44
	0.3	0.65 (0.07)	0.58 (0.07)	0.01 (0.03)	5.49	4.98	0.72
	0.4	0.97 (0.09)	0.81 (0.08)	0.06 (0.04)	9.01	7.74	1.37
225	0.1	0.04 (0.02)	0.03 (0.02)	−0.00 (0.01)	0.24	0.34	0.14
	0.2	0.29 (0.04)	0.27 (0.05)	0.02 (0.02)	2.08	2.14	0.45
	0.3	0.65 (0.07)	0.59 (0.07)	0.01 (0.03)	5.45	5.50	0.79
	0.4	0.97 (0.09)	0.82 (0.09)	0.01 (0.03)	9.24	8.18	1.11
400	0.1	0.03 (0.01)	0.02 (0.02)	−0.00 (0.01)	0.18	0.26	0.11
	0.2	0.22 (0.04)	0.20 (0.04)	0.01 (0.02)	1.56	1.84	0.34
	0.3	0.56 (0.07)	0.51 (0.07)	0.00 (0.03)	4.73	5.08	0.67
	0.4	0.47 (0.08)	0.38 (0.08)	0.03 (0.03)	6.15	6.19	1.12
		Bias for ratio			MSE for ratio		
100	0.1	1.26 (0.18)	0.98 (0.17)	0.66 (0.16)	34.8	29.2	26.7
	0.2	0.59 (0.08)	0.36 (0.07)	0.18 (0.06)	7.12	4.69	3.31
	0.3	0.50 (0.06)	0.24 (0.04)	0.11 (0.04)	3.28	1.80	1.27
	0.4	0.47 (0.04)	0.18 (0.03)	0.05 (0.03)	1.92	0.94	0.65
225	0.1	0.95 (0.12)	0.56 (0.09)	0.15 (0.07)	15.8	8.12	5.49
	0.2	0.56 (0.06)	0.21 (0.04)	0.07 (0.03)	4.37	1.48	1.09
	0.3	0.51 (0.05)	0.15 (0.02)	0.05 (0.02)	2.36	0.63	0.47
	0.4	0.49 (0.04)	0.11 (0.02)	0.04 (0.02)	1.58	0.33	0.25
400	0.1	0.78 (0.11)	0.28 (0.06)	−0.01 (0.05)	12.6	3.71	2.46
	0.2	0.59 (0.07)	0.12 (0.03)	0.00 (0.02)	4.61	0.75	0.54
	0.3	0.55 (0.05)	0.08 (0.02)	0.01 (0.02)	2.45	0.31	0.23
	0.4	0.50 (0.03)	0.09 (0.01)	0.00 (0.01)	1.47	0.18	0.13

Note. Observed standard errors of biases are given in parentheses.

weighted least squares estimator with respect to both bias and MSE, but vice versa for the estimators of the ratio. Results for the exponential model with nugget (Table 4) resembled those for the the exponential model without nugget, in that: (1) the hybrid estimators of partial sill and nugget had slightly smaller bias than the corresponding weighted least squares estimators, but neither estimator was uniformly better with respect to MSE; (2) the hybrid estimator of ratio performed better than its weighted least squares counterpart with respect to both bias and MSE (though this relative superiority was not as substantial as it was for the first two models).

All of these results are consistent with expectation, as there is a theoretical basis for the superiority of the hybrid estimator of the ratio (at least in the case of

Table 2. Empirical Bias and MSE of Semivariogram Parameter Estimates for Spherical Model without Nugget

		Bias for sill		MSE for sill	
n	α_S	WLS	Hyb	WLS	Hyb
100	0.2	−0.08 (0.01)	−0.05 (0.01)	0.08	0.13
	0.4	−0.05 (0.02)	−0.04 (0.02)	0.56	0.43
	0.6	0.55 (0.07)	0.40 (0.06)	5.53	3.69
	0.8	1.28 (0.11)	0.95 (0.09)	13.25	8.84
225	0.2	−0.10 (0.01)	−0.05 (0.01)	0.08	0.10
	0.4	0.00 (0.03)	0.02 (0.02)	0.65	0.41
	0.6	0.73 (0.08)	0.53 (0.07)	7.52	4.66
	0.8	1.30 (0.11)	0.95 (0.09)	14.20	9.07
400	0.2	−0.06 (0.01)	0.01 (0.01)	0.07	0.09
	0.4	0.00 (0.02)	0.04 (0.02)	0.59	0.42
	0.6	0.46 (0.07)	0.33 (0.06)	4.79	3.20
	0.8	1.12 (0.10)	0.84 (0.08)	12.16	7.85
		Bias for ratio		MSE for ratio	
100	0.2	−0.54 (0.04)	−0.48 (0.04)	2.04	1.55
	0.4	−0.11 (0.03)	−0.07 (0.02)	1.06	0.53
	0.6	0.04 (0.03)	−0.02 (0.02)	0.78	0.24
	0.8	0.05 (0.02)	−0.02 (0.01)	0.56	0.13
225	0.2	−0.33 (0.04)	−0.12 (0.03)	1.35	0.80
	0.4	−0.02 (0.03)	0.06 (0.02)	1.02	0.24
	0.6	0.06 (0.03)	0.02 (0.01)	0.83	0.10
	0.8	0.06 (0.02)	0.01 (0.01)	0.58	0.05
400	0.2	−0.25 (0.03)	0.02 (0.02)	1.00	0.47
	0.4	−0.05 (0.03)	0.07 (0.01)	0.84	0.14
	0.6	0.06 (0.03)	0.02 (0.01)	0.71	0.06
	0.8	0.03 (0.02)	0.00 (0.01)	0.45	0.03

Note. Observed standard errors of biases are given in parentheses.

the exponential model without nugget), but no such basis for a superiority of the hybrid estimator of the sill.

In addition to estimation performance, the prediction performance corresponding to each estimation method was evaluated by using the parameter estimates to obtain standard plug-in ordinary kriging variances at sites on a fine grid G, viz. $(0.05, 0.0625, 0.075, \ldots, 0.50) \times (0.05, 0.0625, 0.075, \ldots, 0.50)$ excluding sampling sites. Note that the prediction sites are all located in the lower-left quadrant of the unit square, which is sufficient for our purposes due to isotropy of the models and quadrilateral symmetry of the grid. Prediction performance was measured by ARB, the average (over all sites in G) of the estimated relative bias

Table 3. Empirical Bias and MSE of Semivariogram Parameter Estimates for Rational Quadratic Model without Nugget

		Bias for sill		MSE for sill	
n	α_R	WLS	Hyb	WLS	Hyb
100	0.04	0.02 (0.01)	0.01 (0.01)	0.11	0.12
	0.16	0.03 (0.02)	0.10 (0.03)	0.25	0.72
	0.36	0.03 (0.02)	0.27 (0.04)	0.42	1.88
	0.64	−0.15 (0.02)	−0.16 (0.03)	0.40	0.90
225	0.04	0.01 (0.01)	0.03 (0.01)	0.08	0.12
	0.16	−0.05 (0.01)	0.03 (0.02)	0.16	0.46
	0.36	−0.25 (0.01)	−0.49 (0.02)	0.25	0.62
	0.64	−0.37 (0.02)	−0.77 (0.02)	0.37	1.02
400	0.04	0.00 (0.01)	0.04 (0.01)	0.06	0.16
	0.16	−0.04 (0.01)	0.06 (0.03)	0.15	0.85
	0.36	−0.26 (0.01)	−0.67 (0.02)	0.23	0.83
	0.64	−0.32 (0.02)	−0.90 (0.02)	0.35	1.30
		Bias for ratio		MSE for ratio	
100	0.04	49.4 (6.17)	48.7 (6.15)	40435	40192
	0.16	2.84 (0.38)	2.70 (0.35)	150.6	130.7
	0.36	0.70 (0.07)	0.53 (0.05)	5.21	2.90
	0.64	0.46 (0.03)	0.16 (0.02)	1.42	0.48
225	0.04	6.95 (0.71)	6.67 (0.66)	546.4	476.6
	0.16	0.49 (0.08)	0.36 (0.04)	7.35	1.45
	0.36	0.89 (0.05)	−0.37 (0.02)	3.26	0.70
	0.64	0.48 (0.03)	−0.27 (0.03)	1.06	0.78
400	0.04	3.20 (0.42)	2.98 (0.32)	189.8	112.5
	0.16	1.11 (0.10)	0.57 (0.06)	11.26	4.05
	0.36	0.95 (0.05)	−0.49 (0.04)	3.29	1.88
	0.64	0.58 (0.03)	−0.29 (0.03)	1.21	1.06

Note. Observed standard errors of biases are given in parentheses.

of the plug-in ordinary kriging variance. More precisely, ARB was computed as

$$\frac{1}{|G|n}\sum_{\mathbf{s}\in G}\sum_{i=1}^{n}[M(\mathbf{s};\hat{\boldsymbol{\theta}}(i)) - M(\mathbf{s};\boldsymbol{\theta})]/M(\mathbf{s};\boldsymbol{\theta})$$

where $|G|$ is the number of sites in G, $\hat{\boldsymbol{\theta}}(i)$ is the estimate of $\boldsymbol{\theta}$ for the ith realization, and $M(\mathbf{s};\boldsymbol{\theta})$ is the ordinary kriging variance at site $\mathbf{s}$ corresponding to parameter value $\boldsymbol{\theta}$. ARB combines a measure of performance of the plug-in kriging variance as an estimate of the actual variance of the plug-in kriging predictor, with a measure of performance of the plug-in kriging predictor as an estimate of the ordinary kriging predictor. This is evident upon noting that for a given site $\mathbf{s}$, the

Table 4. Empirical Bias and MSE of Semivariogram Parameter Estimates for Exponential Model with Nugget

n	δ	Bias for partial sill		MSE for partial sill	
		WLS	Hyb	WLS	Hyb
100	0.5	3.43 (0.27)	3.85 (0.44)	83.2	212
	1.0	4.27 (0.29)	3.85 (0.37)	104	149
	2.0	6.13 (0.33)	4.20 (0.29)	148	100
		Bias for nugget		MSE for nugget	
100	0.5	−0.01 (0.01)	0.03 (0.01)	0.19	0.16
	1.0	−0.15 (0.02)	−0.05 (0.02)	0.41	0.36
	2.0	−0.35 (0.03)	−0.18 (0.03)	1.14	0.99
		Bias for ratio		MSE for ratio	
100	0.5	1.50 (0.25)	0.72 (0.24)	64.3	57.7
	1.0	4.16 (0.37)	2.79 (0.35)	157	134
	2.0	9.03 (0.68)	6.63 (0.64)	545	449

Note. Observed standard errors of biases are given in parentheses.

average $(1/n)\sum_{i=1}^{n}[M(\mathbf{s};\hat{\boldsymbol{\theta}}(i)) - M(\mathbf{s};\boldsymbol{\theta})]$ is an estimate of

$$E[M(\mathbf{s};\hat{\boldsymbol{\theta}}) - M(\mathbf{s};\boldsymbol{\theta})],$$

which can be re-expressed (see Stein, 1999, p. 201) as

$$E[M(\mathbf{s};\hat{\boldsymbol{\theta}}) - e(\mathbf{s};\hat{\boldsymbol{\theta}})^2] + E[e(\mathbf{s};\hat{\boldsymbol{\theta}}) - e(\mathbf{s};\boldsymbol{\theta})]^2$$

where $e(\mathbf{s};\boldsymbol{\theta})$ is the prediction error of the ordinary kriging predictor.

Table 5 gives values of ARB for the various models. In the main, the results were similar to the results for estimating the sill-to-range ratio from Tables 1–4. In particular, prediction performance was usually better for the hybrid estimator than for the weighted least squares estimator.

Finally, we compare the amounts of time required to compute the weighted least squares, hybrid, and maximum likelihood estimates for the exponential model with nugget (Table 6). The computations were carried out on a PC running Windows XP and having a Pentium 4 processor of CPU 3.00 GHz and 1.00 GB of RAM. Only results corresponding to a range parameter of 0.4 are displayed in Table 6, as the range seemed to have no effect on computing time. The results indicate that the hybrid method required only marginally (3–7%) more time than weighted least squares, but maximum likelihood estimation was much slower, taking twice as long when $n = 100$ and 30 times as long when $n = 400$ compared to weighted least squares. Also, as the sample size increased from 100 to 400, the

Table 5. Predictor Performance Corresponding to Various Semivariogram Parameter Estimation Methods

n	α_E	ARB ($\times 10^{-4}$)		
		WLS	Hyb	ML
(a) Exponential model without nugget				
100	0.1	105	53	11
	0.2	162	74	17
	0.3	249	94	22
	0.4	350	97	0
225	0.1	149	76	2
	0.2	209	68	10
	0.3	307	79	18
	0.4	401	75	16
400	0.1	134	38	−18
	0.2	229	40	−10
	0.3	341	42	−7
	0.4	431	72	−7
		ARB ($\times 10^{-4}$)		
	α_S	WLS	Hyb	
(b) Spherical model without nugget				
100	0.2	−256	−223	
	0.4	−153	−89	
	0.6	−54	−8	
	0.8	−61	−5	
225	0.2	−177	−62	
	0.4	−70	51	
	0.6	17	6	
	0.8	24	0	
400	0.2	−131	11	
	0.4	−85	63	
	0.6	25	21	
	0.8	−19	0	
		ARB ($\times 10^{-3}$)		
	α_S	WLS	Hyb	
(c) Rational quadratic model without nugget				
100	0.04	−26	−21	
	0.16	44	41	
	0.36	94	81	
	0.64	283	204	
225	0.04	25	25	
	0.16	52	38	
	0.36	447	264	
	0.64	878	460	
400	0.04	19	18	
	0.16	153	96	
	0.36	724	368	
	0.64	1504	680	

Table 5. Continued

n	α_E	ARB ($\times 10^{-4}$) WLS	Hyb
	δ	ARB ($\times 10^{-4}$) WLS	Hyb
(d) Exponential model with nugget			
100	0.5	−165	−18
	1.0	−362	86
	2.0	−398	−18

Note. ARB denotes average relative bias.

average run time for weighted least squares and hybrid estimation increased by less than 50%, while the run time of maximum likelihood estimation increased by almost 2000%.

CONCLUSIONS

In this article we have proposed a method for semivariogram estimation, called hybrid estimation, that combines aspects of weighted least squares and maximum likelihood estimation. What motivated us to consider this method were the infill asymptotic results of Zhang (2004) on consistent estimation in geostatistical models, and empirical evidence that these asymptotic results accurately predict the behavior of estimators obtained from finite samples of moderate size. In particular, these results imply that for the exponential semivariogram, the hybrid estimator of the ratio of sill to range parameter is consistent. Our simulation study of performance indicates that this consistency property, which is shared by the maximum likelihood estimator but not by the weighted least squares estimator, confers better finite-sample properties upon the hybrid estimator of the sill-to-range ratio, and sometimes upon hybrid estimators of other parameters as well, relative to the corresponding weighted least squares estimators (though neither

Table 6. Average Computing Times (in Seconds) for Various Semivariogram Estimation Methods, for the Exponential Model Without Nugget

n	WLS	Hyb	ML
100	0.1205 (0.0006)	0.1237 (0.0006)	0.2449 (0.0018)
225	0.1533 (0.0010)	0.1575 (0.0010)	1.2034 (0.0059)
400	0.1737 (0.0009)	0.1834 (0.0010)	5.1295 (0.0245)

Note. Observed standard errors of computing times are given in parentheses.

performs as well as the maximum likelihood estimators). Moreover, the same consistency property appears to also confer much better finite-sample properties upon plug-in predictors using hybrid estimates when compared to plug-in predictors using weighted least squares estimates. Finally, our investigation suggests that the relative superiority of hybrid estimation to weighted least squares estimation for the sill-to-range ratio and for plug-in prediction extends to models with semivariograms other than the exponential, including some for which no analogous consistency results are yet known.

Although maximum likelihood estimation has better theoretical properties than weighted least squares estimation and has therefore been the recommended method for semivariogram estimation provided the data set is not so large as to make it computationally infeasible, it is presently not supported by much of the menu-based geostatistical software available to geo-scientists. Consequently, many such scientists estimate semivariograms using weighted least squares, which is supported by widely available software. Because hybrid estimation appears to outperform weighted least squares estimation for some models and parameters and seems to always yield substantially better plug-in predictors than weighted least squares estimation, yet requires negligible additional computing effort, we recommend that it supplement, if not replace, weighted least squares as the estimation method of choice among those geostatistical practitioners who do not use maximum likelihood. Accordingly, we suggest that software that currently estimates the semivariogram by weighted least squares methods be amended to include hybrid estimation as an option.

ACKNOWLEDGMENTS

We thank the editors and the referees for their constructive comments, which led to improvements in the article. The research of Zhang is partially supported by the National Science Foundation of the USA (DMS-0405782).

REFERENCES

Clark, I., 1979, Practical geostatistics: Applied Science Publishers: Essex, England, 129 p.

Cressie, N., 1985, Fitting variogram models by weighted least squares: Math. Geol., v. 17, no. 5, p. 563–586.

Curriero, F. C., and Lele, S., 1999, A composite likelihood approach to semivariogram estimation: Journal of Agricultural, Biological, and Environmental Stat., v. 4, no. 1, p. 9–28.

Dubin, R. A., 1994, Estimating correlograms: A Monte Carlo study: Geographical Systems, v. 1, p. 189–202.

Ecker, M. D., and Gelfand, A. E., 1997, Bayesian variogram modeling for an isotropic spatial process: Journ. of Agricultural, Biological, and Environmental Stat., v. 2, no. 4, p. 347–369.

Ecker, M. D., and Gelfand, A. E., 1999, Bayesian modeling and inference for geometrically anisotropic spatial data: Math. Geol., v. 31, no. 1, p. 67–83.

Genton, M. G., 1998, Variogram fitting by generalized least squares using an explicit formula for the covariance structure: Math. Geol., v. 30, no. 4, p. 323–345.
Handcock, M. S., and Stein, M. L., 1993, A Bayesian analysis of kriging: Technometrics, v. 35, no. 4, p. 403–410.
Journel, A. G., and Huijbregts, C. J., 1978, Mining geostatistics: Academic Press, London, 600 p.
Kitanidis, P. K., 1983, Statistical estimation of polynomial generalized covariance functions and hydrologic applications: Water Resources Res., v. 19, no. 4, p. 909–921.
Mardia, K. V., and Marshall, R. J., 1984, Maximum likelihood estimation of models for residual covariance in spatial regression: Biometrika, v. 71, no. 1, p. 135–146.
Stein, M. L., 1988, Asymptotically efficient spatial interpolation with a misspecified covariance function: Annals of Stat., v. 16, no. 1, p. 55–63.
Stein, M. L., 1990, Uniform asymptotic optimality of linear predictions of a random field using an incorrect second-order structure: Annals of Stat., v. 18, no. 2, p. 850–872.
Stein, M. L., and Handcock, M. S., 1989, Some aymptotic properties of kriging when the covariance function is misspecified: Math. Geol., v. 21, no. 2, p. 171–190.
Stein, M. L., 1999, Interpolation of spatial data: Springer, New York, 247 p.
Ying, Z., 1991, Asymptotic properties of a maximum likelihood estimator with data from a Gaussian process: Jour. of Multivariate Analysis, v. 36, no. 2, p. 280–296.
Zhang, H., 2004, Inconsistent estimation and asymptotically equal interpolations in model-based geostatistics: Jour. of the Am. Stat. Association, v. 99, no. 465, p. 250–261.
Zimmerman, D. L., 1989, Computationally efficient restricted maximum likelihood estimation of generalized covariance functions: Math. Geol., v. 21, no. 7, p. 655–672.
Zimmerman, D. L., and Zimmerman, M. B., 1991, A comparison of spatial semivariogram estimators and corresponding ordinary kriging predictors: Technometrics, v. 33, no. 1, p. 77–91.