APPLICATION OF BIASED ESTIMATION
TO CLOSE-RANGE PHOTOGRAMMETRIC DATA REDUCTION

Y.D. Huang
Photogrammetric Laboratory, Civil Engineering Department
Northern Jiaotong University, Beijing, P.R. China

ABSTRACT

The theory of biased estimation in statistics is introduced. The application of the theory to photogrammetric data reduction is discussed. The weighted least squares methods with supplemented fictitious weighted direct observations of additional parameters are found to conform to the principle of biased estimation. Two of these methods are investigated with experiments in the close-range cases. It is shown that the two methods can be more accurate, at least safer than both the least squares self-calibrating method with free additional parameters and the non-self-calibrating method without additional parameters.

INTRODUCTION

Self-calibration is now widely applied in photogrammetric practice. While it has been found not to act effectively in all cases. Sometimes it even deteriorates the final results. That is because self-calibration always enlarges the variances of the least squares estimators at the same time as it rectifies their biases. If the biases that have been rectified cannot compensate for the gains of the variances, deteriorated results will appear. So it is evident that in order to make self-calibration more effective both variance and bias must be considered and their resultant effect minimized. To do so, an important way is to improve the estimating methods. The theory of biased estimation of the coefficients in linear regression has developed a kind of estimating methods to improve the least squares method. The philosophy of the theory is to decline variance in a beneficial deal at the sacrifice of moderate increase of bias so as to acquire a minimum mean square error in the final result. The theory developed rapidly in last decade or more in applied statistics, and has come into application in some fields. In the following, the theory of biased estimation will be introduced and its relationship with photogrammetric data reduction be discussed so that the results of studies in biased estimation will be applied to photogrammetry. Two estimating methods will be constructed for self-calibration according to the principle of biased estimation, and then examined with simulated experiments in close-range photogrammetry.

FUNDAMENTAL PRINCIPLE OF BIASED ESTIMATION

As is known from statistics, a least squares estimator is a minimum variance unbiased estimator (for the cases with normal distribution). However, in some cases, especially in self-calibration, the minimum variances that the least squares estimators possess are still unacceptably large probably since strong correlations and ill-conditioned coefficient matrices are present. Thus, some improvement of the least squares method is necessary in these cases.

Because the least squares estimator is the best of unbiased estimators, to improve the least squares estimator, we must resort to biased estimators and seek the improvement in the sense of mean square error. In applied statistics some kinds of biased estimators were developed for the purpose. The ridge estimators and the generalized ridge estimators are among them.

For the linear model

$$L = AX + V_r \quad , \tag{1}$$

the ridge estimator of X, marked by $\tilde{X}$, is defined by the following formula

$$\tilde{X}(k) = (N + kI) A'L \tag{2}$$

where $N=A'A$ and $0 \leq k < \infty$, $L:(n*1)$, $X:(m*1)$.
The bias, variance and mean square error of $\tilde{X}(k)$ are as follows:

$$Bias(\tilde{X}(k)) = E(\tilde{X}(k)) - X = ((I + kN^{-1})^{-1} - I) X \quad , \tag{3}$$

$$Var(\tilde{X}(k)) = \sigma^2 (N + kI)^{-1} N (N + kI)^{-1} \quad , \tag{4}$$

$$\begin{aligned} Mse(\tilde{X}(k)) &= E((\tilde{X}(k) - X)'(\tilde{X}(k) - X)) \\ &= tr(Var(\tilde{X})) + (Bias(\tilde{X}))'(Bias(\tilde{X})) \quad . \end{aligned} \tag{5}$$

It has been proved that the variance term of the last formula is a monotone decreasing function of k , and that the bias term is a monotone increasing function of k . It follows that
if k varies from 0 to ∞ ,
then $tr(Var(\tilde{X}))$ varies from $\sigma\, tr(N^{-1})$ to 0 and
$(Bias(\tilde{X}))'(Bias(\tilde{X}))$ from 0 to $X'X$.
It has also been proved that if $X'X$ has boundary, then there exists some $k_0>0$ which makes the ridge estimator $\tilde{X}(k)$ possess a smaller mean square error than the least squares estimator $\hat{X}$, that is

$$Mse(\tilde{X}(k_0)) < Mse(\tilde{X}(0)) = Mse(\hat{X}) \quad . \tag{6}$$

This property shows that properly choosing k can make the ridge estimator better **in** the **sense** of mean square error than the least squares estimator.

Assuming λ_i to be the ith eigenvalue of N , we obtain a expanded form of formula (5) in which the relation between k and the $Mse(\tilde{X}(k))$ is clearer.

$$Mse(\tilde{X}(k)) = \sigma^2 \sum_{i=1}^{m} \frac{\lambda_i}{(\lambda_i + k)^2} + k^2 X' (N + kI)^{-2} X$$

It can be seen from above that a reasonable choice of k depends on the true values of the unknowns σ and X . It is evidently impossible to find some k always making $Mse(\tilde{X}(k))$ minimum or smaller than $Mse(\hat{X})$ whatever true values σ and X take. So, in practice, we can only determine k according to experience or sample of observations containing the information of σ and X . Up to now quite a few principles and methods of choosing k have been developed in applied statistics. Who is intereasted can refer to the listed literature (Hocking,1976; Wichern, 1978). But the effectiveness of each of these methods is not uniform, varying with the coefficient matrice and the true values of σ and X . None of them is better than the others in all cases. Therefore, special procedures of choosing k aimed at special cases need to be developed to apply ridge estimation effectively. These remarks also apply to the generalized ridge estimator.

The generalized ridge estimator can be defined as follows:

$$\tilde{X}_g(K) = (N + Z'KZ)^{-1} A'L \tag{8}$$

where $K = diag(k_1, k_2, \cdots, k_m)$ is a diagonal matrix in which the diagonal elements may be different, and Z is the orthogonal matrix which makes

$$ZNZ' = \Lambda = diag(\lambda_1, \lambda_2, \cdots, \lambda_m) \quad .$$

Let $Y=ZX$, then

$$\tilde{Y}_g(K) = (\Lambda + K)^{-1} ZA'L \tag{9}$$

is the generalized estimator of Y . It can be found that the generalized ridge estimators have included the (simple) ridge estimator as their special case with all the elements of K being identical. They can decline the mean square errors further. In fact,

$$Mse(\tilde{X}_g(K)) = Mse(\tilde{Y}_g(K)) = \sigma^2 \sum_{i=1}^{m} \frac{\lambda_i}{(\lambda_i + k_i)^2} + \frac{k_i^2\, y_i^2}{(\lambda_i + k_i)^2} \tag{10}$$

Derivate the right side of above formula to each k_i , let it equal to zero and solve it. Then it follows that

$$k_i = \sigma^2 / y_i^2 \qquad (i = 1,2,\cdots,m) . \qquad (11)$$

These k_i will make the $Mse(\tilde{X}_g(K))$ and $Mse(\tilde{Y}_g(K))$ minimum. Generally one y_i is different from another, so the best k_i given by formula(11) are not the same. Therefore, theoretically, generalized ridge estimators can reach smaller mean square errors than simple ridge estimators.

On ridge estimation (simple and generalized), the following remarks can also be given , which will be helpful for us to understand this new estimation.

When k or K is independent of the sample, the ridge estimator is the linear function of the sample, and called linear estimator. Because in application k or K is always determined from the sample, the ridge estimator in practice essentially belongs to nonlinear estimators. It is known from statistics that if there exists some estimator uniformly better than the least squares estimator, it must be nonlinear.

The ridge estimator $\tilde{X}(K)$ is equivalent to the weighted least squares estimator obtained by supplementing unknowns' fictitious direct observations of values 0 and variance matrix $\sigma^2 K^{-1}$. In fact, in so doing, the augmented error equations become

$$\begin{pmatrix} L \\ 0 \end{pmatrix} = \begin{pmatrix} A \\ I \end{pmatrix} X + \begin{pmatrix} V_1 \\ V_x \end{pmatrix} \qquad \text{and weight matrix:} \begin{pmatrix} I & 0 \\ 0 & K \end{pmatrix}$$

the normal equations become

$$(A'A + K) X^* = A'L \quad .$$

Then $X^* = \tilde{X}(K)$. $\tilde{X}(K)$ can also be acquired by using the constrained least square method in which the constraint is $X'X \leq d^2$, where d is related to K . If some priori statistical information like $X \sim N(0, \sigma^2 K^{-1})$ is assumed, $\tilde{X}(K)$ can again be obtained by applying the maximum posteriori estimation or Bayess principle. The three different ways lead to the same ridge estimator. However they have all in all different meanings, systems of research and results of study.

When k_i is the infinite, the ridge estimation is equivalent to deleting the unknown corresponding to the k_i from the model, taking zero as its estimator and then conducting least squares estimation. And when k_i is zero, the ridge estimator of the corresponding unknown degenerates into the least squares estimator (for y_i). These show that ridge estimation can play the role of selecting parameters, furthermore can treat the two decisions of accepting or rejecting parameters as a continuous process and make a compromise. Therefore, theoretically, ridge estimation is more capable of improving the quality of the estimators of unknowns than parameter selection approaches.

Unfortunately, none of the ridge estimators proposed so far can improve the least squares estimator uniformly. Where a ridge estimator can be better than the least squares estimator is a complicated problem on which the present theoretic research work is not yet satisfactory. But it should be realized that an important aspect of the application of ridge estimation lies in the ingenious combination of the principles with the concrete practices.

The application of self-calibration has brought photogrammetric data reduction a series of new problems, such as the determination of error models, the selection of parameters and the treatment of strong correlation. These problems essentially belong among those how to consider both the variance and the bias so as to reach the best actual quality, for which the biased estimation was originally proposed. Now the theory of biased estimation has become a special theory, developed a series of methods and come into application in some fields. Therefore, it can be expected that applying the principle and the methods of biased estimation to photogrammetric data reduction will be beneficial.

ON THE APPLICATION OF BIASED ESTIMATION TO PHOTOGRAMMETRIC DATA REDUCTION

The theory of biased estimation was developed for linear models. While the models encountered in photogrammetry are nonlinear. It is therefore needed to deal with the problems of the relation and transference between linear models and nonlinear models before applying the biased estimation to photogrammetry. For nonlinear models, the usual principle for the estimation of their unknowns is also the least squares principle. That is, for the following nonlinear model

$$L = F(X) + V_{\mathbf{r}} \quad , \tag{14}$$

to find the $\hat{X}$ as the estimator of X making

$$(L - F(\hat{X}))'(L - F(\hat{X})) = \min \quad . \tag{15}$$

where $F = (f_1, f_2, \cdots, f_n)'$, every f_i $(i=1,\cdots,n)$ is the nonlinear function of X. The usual method for acquiring this estimator is the Gauss-Newton method which is currently used in photogrammetry. As we know, it is actually the recursive application of the linear least squares method. It is this link between Gauss-Newton method and linear least squares method that reference will be made to in the following.

Although nonlinear least squares estimators are widely used in practice, it is not yet clear how their statistical properties are. No general conclusion has been drawn about that. Generally speaking, the least squares estimators of nonlinear parameters do not necessarily possess such properties as being unbiased or efficient or so like linear least squares estimators. Their statistical properties vary with various models. Because the solution of nonlinear least squares estimation is an iterative and approximate one, to investigate its statistical properties is quite difficult. For this reason, when using the nonlinear least squares estimators, people always consider the problems relevant to the statistical properties approximately as if they were with the linear cases(see Afifi,1979). This practice has become conventional and granted. In the consideration of applying the biased estimation to the nonlinear photogrammetric problems, the similar approximation may be made. Stress may be put on proposing new estimating methods and testing them by experiments.

In model (14), if the nonlinear least squares method is adopted to estimate X, and the initial approximate values of $\hat{X}$, marked by X_0 , is sufficiently close to their least squares solutions, then the estimator can be expressed approximately as the linear function of the observation L as follows:

$$\hat{X} = (A_0'A_0)^{-1}A_0'(L - F(X_0)) + X_0 \tag{16}$$

in which

$$A_0 = \frac{\partial F}{\partial X}\bigg|_{X=X_0} \quad .$$

Assume $V_{\mathbf{r}}$ in model (14) to follow $N(0,\sigma^2 I)$, then the variance of $\hat{X}$ will approximately be

$$\mathrm{Var}(\hat{X}) = \sigma^2(A_0'A_0)^{-1} \tag{17}$$

(this is sometimes very coarse). This shows that the variance of X has much to do with the condition of matrix A_0 . If there is some strong correlation in A_0 , the eigenvalues of $A_0'A_0$ will be very small, and that enlarges the $tr(Var(\hat{X}))$ severely. Therefore, according to the thoughts of biased estimation, a natural way to improve the least squares estimator is to take the estimator of the following form so as to shrink the variance of the estimator.

$$\tilde{X}(K) = (A_0'A_0 + K)^{-1}A_0'(L - F(X_0)) + X_0 \tag{18}$$

If $\hat{X}$ is assumed to be unbiased, then $\tilde{X}$ is biased. It is thus needed to choose some proper K to minimize the mean square error of the estimator as in linear cases. So far it can be seen that if it has been assumed that the nonlinear least squares estimator is unbiased, its variance as formula (17) and X_0 close enough, then the problem of improving nonlinear least squares estimators has no difference from the same problem in linear cases.

Many results of theoretic studies on biased estimation are with the transformated unknowns Y as in formula (9), not the original X . However in photogrammetric practice, such transformation is of great amount of computation. Therefore it is still difficult at present to apply those results of theoretic studies directly. But some alteration may be made, such as, to use the following biased estimator instead of that given by formula (8):

$$\tilde{X}(K) = (N + K)^{-1} A' L \quad . \qquad (19)$$

This estimator can also possess smaller variance than the least squares estimator.

THE METHODS, EXPERIMENTS AND RESULTS

It is known from above two sections that the nonlinear weighted least squares methods with supplemented fictitious weighted zero observations of the unknowns conform to the principle of biased estimation. So this kind of methods can be used to improve the least squares method in photogrammetry. In this study two methods of this kind were constructed for self-calibration in close-range photogrammetry and examined by experiments. They are featured below by their weight functions, or biasing terms.

For the Method 1 the weight function is

$$p_{s_i}^{k+1} = \sigma^2 (1 - p_{s_i}^{k} q_{s_i}^{k}) / s_i^k \quad \text{with} \quad p_{s_i}^{1} = 0 \ , \ s_i^o = 0 \quad (i=1,\cdots,m_s) \qquad (20)$$

where σ is the first estimator of the reference variance,
k_k is the ordinal number of estimation,
s_i^k is the kth estimator of s_i, the additional parameter,
p_i^k is the weight for the kth estimation,
q_i^k is the diagonal element of $(A_k' A_k + P_x)^{-1}$ corresponding to s_i , in which

A_k is calculated with the kth estimators, P_x is the diagonal weight matrix of unkowns in which the diagonal elements except those corresponding to the additional parameters s are zero all the time. This method is executed by recursively determining the weights by formula(20) and repeatedly conducting the nonlinear weighted least squares adjustment with the current weights each time until some k for which each s_i^k has little difference from its last value s_i^{k-1} . The Method2 differs from the Method1 only in that it takes the following weight function:

$$p_{s_i}^{k+1} = p_s^{k+1} = \sigma^2 \sum_{j=1}^{m_s} (1 - p_s^k q_{s_j}^k) / \sum_{j=1}^{m_s} (s_j^k)^2 \qquad (i=1,\cdots,m_s) \ , \qquad (21)$$

where the symbols have the same meanings as those in the Method1. As another alteration, the two methods apply biasing terms only to the additional parameter part of the unknowns. That is based on the consideration that the range of values of additional parameters in selfcalibration is very limited to which the methods can adapt themselves, and that the extent of approximation of the initial values of the main unknowns is changeable for the methods of producing the initial values are full of variety, so that a method with applying biasing terms to the main unknowns could lose its adaptability. The similar methods were once proposed and used in aerotriangulation before. But special studies on them and experiments aimed at concrete close-range cases are still not sufficient.

To examine the effectiveness of the two methods in close-range photogrammetry, a set of simulated experiments were specially made. They are described as follows. All are simulated by a computer.

Cameras with f=58 mm and formats being 24*36 mm^2 are used for photographing.

Stereo pairs of photographs are taken with nearly 100% overlap and 20% depth of field. Eighty object points are normally distributed within the drawn field. The control data are provided by designating certain number of true coordinates values from the object points. The number has three options of 21, 12 and 6. The observations of image coordinates are obtained by damaging the true image coordinates with artificial errors. The random errors follow the normal distribution $N(0, \sigma^2)$, in which σ has the choice of 1µ, 3µ, 9µ and 22µ. The systematic errors are produced according to the classical physical model:

$$x_s = x_o + (x-x_o)(k_1 r^2 + k_2 r^4 + k_3 r^6) + \\ + p_1(r^2 + 2(x-x_o)^2) + 2p_2(x-x_o)(y-y_o) + A(y-y_o) \quad , \tag{22a}$$

$$y_s = y_o + (y-y_o)(k_1 r^2 + k_2 r^4 + k_3 r^6) + \\ + p_2(r^2 + 2(y-y_o)^2) + 2p_1(x-x_o)(y-y_o) + B(y-y_o) \quad , \tag{22b}$$

in which x and y are image coordinates, $r^2 = (x-x_o)^2 + (y-y_o)^2$ and the other symbols are parameters which are given three sets of values in setting conditions and taken as additional unknown parameters in adjustments.

code name	position of photo	x_o 10^{-2}mm	y_o 10^{-2}mm	k_1 $10^{-6}mm^{-2}$	k_2 $10^{-8}mm^{-4}$	k_3 $10^{-10}mm^{-6}$	p_1 $10^{-5}mm^{-1}$	p_2 $10^{-5}mm^{-1}$	A 10^{-3}	B 10^{-3}
G	L	0.1	0.1	0.3	0.1	0.2	-0.2	0.2	-0.1	0.4
	R	-0.1	-0.1	0.4	-0.1	0.3	0.4	-0.1	0.5	0.1
D	L	10	10	3	1	-1	2	-2	5	-10
	R	5	-10	4	1	-2	1	2	-12	8
H	L	100	100	0.3	0.1	-0.1	0.2	-0.2	50	-10
	R	50	-100	0.4	0.1	-0.2	0.1	0.2	-12	8

Table 1: Three Sets of Simulated Systematic Errors

To compare the two methods with the least squares self-calibrating method with free additional parameters and the non-self-calibrating method without additional parameters, all the four methods have been executed to every case examined. The root mean square of the real errors of position of all object points has been taken as the criterion for comparison.

The results of the experiments are listed in table2--5. It is shown by the results that:

1. Method1 and Method2 can acquire more accurate results than both the least squares self-calibrating method with free additional parameters and the non-self-calibrating method without additional parameters, at least, if not in some cases, still be closer to the better of them. This might be true for any number of additional parameters. These two methods can thus be said to be safe and robust in respect that this effect of the two methods to systematic errors is much like that of the "robust method" to gross errors (see Kubik,1984).
2. In the cases of large random errors and variance factors and small systematic errors, the two methods are much better than the least squares method which is now at a disadvantage. While, in the contrary cases in which the least squares method turns superior, the two methods have little improvement and sometimes are even not as good as the least squares method. That is shown in table2,3and4.
3. When the number of control data decreases, the least squares method becomes worse, and the two methods improve it a lot. This means that using the two methods can save control data. That is shown in table4.
4. The two methods have almost the same trend. The cases in which Method2 is better than Method1 are in the majority. But in the cases where the values of systematic error parameters have great disparity like set H, Method1 is considerably superior to Method2 as shown in table5.

CONCLUDING REMARKS

After introducing the theory of biased estimation, we have found that the theory has close relation with many problems and methods in photogrammetry. The weighted least squares method with supplemented fictitious weighted direct observations of additional parameters just conform to the principle of biased estimation. By simulated experiments in close-range photogrammetry, two methods of this kind have been proved to be more accurate , at least safer than the least squares methods with or without additional parameters. But the minute behaviour of these methods has not yet been discerned and further research in this direction is required. The theory of biased estimation has a wealth of content, only a little part of which has been introduced here. The work of introducing and linking in the paper has been done in the expectation that the theory of biased estimation will play its special directive role in photogrammetric data reduction.

REFERENCES

Afifi,A.A. and Azen,S.P.(1979),"Statistical Analysis, A Computer Oriented Approach",(2nd edi.), Academic Press, New York.

Hoerl,A.E. and Kennard,R.W.(1970),"Ridge Regression: Biased Estimation for Non-orthogonal Problems", Technometrics, 12, 55-68.

Hocking,R.R., Speed,F.M. and Lynn,M.J.(1976),"A Class of Biased Estimators in Linear Regression", Technometrics, 18, 425-438.

Hemmerle,W.J. and Brantle,T.F.(1978),"Explicit and Constrained Generalized Ridge Estimation", Technometrics, 20, 109-120.

Huang,Y.D.(1986),"More Remarks on Self-calibration", Presented Paper of the ISPRS Comm. V Symposium in Ottawa.

Kubik,K.,Weng,W. and Frederiksen,P.(1984),"Oh,Grosserrors!", ISPRS,Comm.III, Rio.

Marquardt,D.W.(1970),"Generalized Inverses, Ridge Regression, Biased Linear Estimation and Nonlinear Estimation", Technometrics, 12, 591-612.

Wichern,D.W. and Churchill,G.A.(1978),"A Comprision of Ridge Estimates", Technometrics, 20, 302-311.

Case	σ_μ	LS,WithoutAP		LS with AP		Method 1			Method 2		
		$RMSV_\mu$	$RMSPE_\mu$	$RMSV_\mu$	$RMSPE_\mu$	$RMSV_\mu$	$RMSPE_\mu$	$RIP_\%$	$RMSV_\mu$	$RMSPE_\mu$	$RIP_\%$
a	1.0	5.23	9.76	1.21	3.59	1.18	2.40	+33	1.13	3.34	+ 7
	3.0	5.99	10.76	3.37	10.42	3.46	7.19	+31	3.42	6.23	+40
	9.0	10.93	19.12	9.98	31.13	10.34	19.55	+37	10.40	18.05	+42
	22.0	24.46	43.30	24.03	104.28	25.47	46.58	+55	25.22	45.45	+56
b	1.0	4.22	12.12	1.23	2.45	1.19	2.27	+ 7	1.19	2.02	+18
	3.0	5.09	12.64	3.47	7.29	3.60	6.28	+14	3.62	5.29	+19
	9.0	10.39	19.68	10.48	21.94	10.85	18.77	+14	10.75	17.68	+19
	22.0	24.14	43.05	25.75	54.65	26.35	44.65	+18	26.28	43.35	+21
c	1.0	4.02	15.65	1.81	6.45	1.22	3.57	+45	1.21	2.49	+62
	3.0	5.04	15.05	4.60	17.90	3.64	7.93	+56	3.62	6.95	+61
	9.0	10.64	20.08	10.38	45.73	10.91	21.42	+53	10.86	20.86	+54
	22.0	24.73	44.58	25.66	101.82	26.81	46.94	+54	26.62	50.03	+51
d	1.0	2.23	22.16	1.13	7.17	1.13	6.25	+13	1.16	5.70	+21
	3.0	4.01	24.02	3.36	21.44	3.48	16.66	+22	3.48	15.71	+27
	9.0	10.46	45.98	10.05	64.06	10.49	50.17	+22	10.48	48.30	+25
	22.0	24.89	109.98	24.59	156.87	25.57	121.03	+23	25.63	120.85	+23

Table 2: Comparison in the Cases with Small Systematic Errors Set G

RMSV: the Square Root of $V'V/(n-m)$, n & m are numbers of observations & unknowns.
RMSPE: the root-mean-square of real position errors.
RIP: the relative improvement percentage to the RMSPE for "LS with AP".

Case $\sigma=9\mu$	Systematic Error Set	LS with AP RMSPE μ	Method 1		Method 2	
			RMSPE μ	RIP %	RMSPE μ	RIP %
a	G	31.13	19.55	+37	18.05	+42
	D	29.99	26.00	+13	23.83	+21
b	G	21.94	18.77	+14	17.68	+19
	D	22.46	20.62	+ 8	21.01	+ 6
c	G	45.73	21.42	+53	20.86	+54
	D	43.65	31.55	+28	26.18	+40
d	G	64.06	50.17	+22	48.30	+25
	D	61.59	50.90	+17	49.23	+20

Table 3: Comparison under Different Systematic Errors (Set G and Set D)

Case		Control Data	LS with AP RMSPE μ	Method 1		Method 2	
	(μ)			RMSPE μ	RIP %	RMSPE μ	RIP %
AP: G	σ= 3.0	21	7.29	6.28	+14	5.92	+19
		12	17.90	7.93	+56	6.95	+61
	σ=22.0	21	54.65	44.65	+18	43.35	+21
		12	101.82	46.94	+54	50.03	+51
AP: D	σ= 3.0	21	7.66	9.78	-28	7.88	- 3
		12	19.02	12.69	+33	9.36	+51
	σ=22.0	21	54.20	47.40	+13	48.71	+10
		12	96.48	79.38	+18	68.34	+29
AP: G, σ= 3.0		21	10.42	7.19	+31	6.23	+40
		12	50.31	12.23	+76	9.92	+80
		6	61.71	8.94	+86	9.04	+85

Table 4: Comparison under Different Numbers of Redundant Control Coordinates

Case	LS with AP RMSPE μ	Method 1		Method 2	
		RMSPE μ	RIP %	RMSPE μ	RIP %
a	3.97	2.74	+31	3.98	- 0.3
b	11.08	8.20	+26	11.15	- 0.6
c	34.56	26.73	+23	35.79	- 4
d	88.53	63.13	+29	66.66	+25
e	72.64	58.00	+20	68.96	+ 5
f	19.34	16.48	+15	18.81	+ 3
g	23.04	14.49	+37	21.66	+ 6

Table 5: Comparison of Method 1 with Method 2 under Set H