HANDLING CONCENTRATIONS BELOW QUANTIFICATION LIMIT IN POPULATION PHARMACOKINETICS

M. TOD

Département de Pharmacotoxicologie, Hôpital Avicenne,
and Laboratoire de Pharmacie Clinique, Faculté de Pharmacie, Paris,
France

INTRODUCTION

In population pharmacokinetics, discarding BQL data may result in a more biased and/or less precise estimation of the population parameters (1,2). Previous studies have proposed several strategies in the context of NONMEM and assessed their relevance in the case of a one compartment PK model (1,2). In this study, three practical strategies for handling BQL values were compared on simulated data with a two-compartment model. All these estimations were made with WinBUGS (3). A limited exploration of the influence of the information content of the design on the performance of each method has been made.

DEFINITION OF SYMBOLS

y_j the vector of observations gained in the j-th individual

$f(.)$ the predicted concentration, t_j the vector of times

P_j. the p-vector of individual pharmacokinetic parameters

ε_j the vector of residual errors with mean 0 and covariance matrix Σ

g_{ij}^2 the elements of Σ *i.e.* the variance attached to the i-th observation in individual j

θ the vector of fixed effects

Ω the variance-covariance matrix of η's

CL elimination clearance, Q distribution clearance, V_1 central volume, V_2 peripheral volume

THEORY

The population model is a three-stage hierarchical model

First stage: residual error model $y_j = f(P_j, t_j) + \varepsilon_j \qquad \varepsilon \sim N(0, \Sigma)$ (1)

Σ is diagonal with elements $g_{ij}^2 = \left[\sigma_1 . f(P_j, t_i) + \sigma_2\right]^2$ (2)

Second stage: interindividual variability $\log P_j \sim N(\log \theta, \Omega)$ (3)

Third stage: priors for the population parameters θ, Ω and σ

$\sigma_1^{-2} \sim Gamma\ (a, b) \qquad \log \theta \sim N(M, C) \qquad \Omega^{-1} \sim Wishart\ (R, p)$ (4)

where *a*, *b*, *M*, *C*, *R* and p are fixed.

The likelihood for the population parameters is the product of the contributions of each individual, L_j:

$$L_j(\theta, \Omega, \sigma^2) = \int p(y_j / \theta, \eta_j, \sigma^2).p(\eta_j / \Omega).d\eta \qquad (5)$$

where

$$p(y_{ij} / \theta, \eta_j, \sigma^2) \propto g_{ij}^{-1} \, exp\left[(y_{ij} - f(P_j, t_{ij}))^2 / g_{ij}^2\right] \quad \text{if } y_{ij} > QL \qquad (6)$$

or $$p(y_{ij} / \theta, \eta_j, \sigma^2) \propto \int_0^{QL} exp\left[(y_{ij} - f(P_j, t_{ij}))^2 / g_{ij}^2\right] dy_{ij} \quad \text{if } y_{ij} < QL \qquad (7)$$

However, most commercially available softwares unless WINBUGS do not implement equation (7), so that BQL values cannot be handled correctly.

METHODS FOR HANDLING BQL VALUES

Method 1 consists in discarding BQL values from the data.

Method 2: the BQL values are fixed to QL/2, while σ_2 is fixed to QL/4. If there are several consecutive BQL values, only the first is kept.

Method 3 or reference method makes use of equation (7). The BQL data are replaced by the missing value code in the data file, and the bounds of the integral of equation (7) must be supplied.

Method 4 or gold standard method is analogous to method 1 but with no BQL values, i.e. all the observations are available.

The performances of the methods were assessed by simulation of pseudoexperimental data. The specific example is as follows.

SIMULATION OF THE DATA (1)

Drug disposition: open bicompartmental model, IV bolus, dose 1000 mg.

Individual parameters were generated according to equation (3). The median of (CL, Q, V_1, V_2) distribution was θ = (3.5, 1, 10 ,10) (typical half-lives of 1.5 and 9.5 h). The interindividual CV of each parameter was 0.25. "Observed" concentrations were generated with $\sigma_1 = 0.15$ and $\sigma_2 = 0$.

Four observations per individual, two sampling schedules:

The population D-optimal schedule (4) T_D = (0.1, 4.5, 15, 32),

The nearly-optimal schedule T_{nonD} = (0.25, 8, 20, 32).

The ratio of the determinant of the population Fisher information matrix calculated for these two schedules is 0.99

SIMULATION OF THE DATA (2)

In a given data set, the same schedule was applied to all individuals. The first three samples were always above QL, the fourth was occasionally BQL.

In some data sets (denoted as $m = 1/3$), samples were discarded so that the number of samples (among the first three) in a given individual could be 0, 1, 2 or 3. In other data sets (denoted as $m = 3/3$), this random elimination procedure was not applied.

Data sets were generated for various combinations of the values of the population size N, the total number of data points n, the proportion m, the quantification limit QL and the sampling times T, as shown in **table I**.

Ten replicates of each data set were obtained by generating samples from the distributions of P and ε. A plot of the data generated for a single replicate is shown in **figure 1**.

ESTIMATION OF THE POPULATION PARAMETERS

The population parameters $\psi = (\theta, \Omega, \sigma_1)$ were estimated for each replicate of all data sets described in table I using WinBUGS 1.3 (2), with non-informative priors. For each replicate, the mean and the SD of the posterior distribution of the 9 elements of ψ was obtained.

MEASURES OF PERFORMANCE

For each replicate, bias and precision were measured by:

the mean absolute error: $$MAE = \frac{100}{9}\sum_{k=1}^{9} Abs\left(\frac{\hat{\psi}_k - \psi_k}{\psi_k}\right)$$

the mean standardized standard error: $$MSSE = \frac{100}{9}\sum_{k=1}^{9}\left(\frac{SE_k}{\psi_k}\right)$$

RESULTS AND DISCUSSION (1)

The two summary statistics MAE and MSSE for the ten replicates of any test case were very reproducible: Hence, only the means of the ten replicates are reported in the bar graphs, not the ranges.

A very similar pattern was observed in all cases: the performances of the methods ranked in the **order 1 < 2 < 3 < 4**. Method 1 performed really poorly in some cases, method 2 achieved a good accuracy but a poor precision, and method 3 achieved in many cases the best possible performances, i.e. those of method 4. However, performances of method 2 may depend on the imputed value (here QL/2), the standard deviation (here QL/4), and the terminal half-life. These points remains to be studied.

RESULTS AND DISCUSSION (2)

The best performances were observed with **case I** (typical MAE and MSSE #10%), *i.e.* a relatively reach data situation (3 or 4 measures per individual, a total of 240 measures), with a small proportion of BQL values (10%).

In case II, a typical sparse data situation (same conditions as case I but a smaller number of samples per individual, a total of 120 measures and a 20% proportion of BQL values), MAE and MSSE rise to #15%. When the proportion of BQL values was increased to 30% by increasing QL (**case IV**), the statistics did not increase further, unless for the method 1. Case II can also be compared to **case III**, where the conditions were the same but the number of individuals was increased 3-fold. This resulted in an improved estimation (reduction of both statistics to about 10%). Hence, the differences between method 1 and the

other methods were increased when the amount of information was lower and when the proportion of BQL values increased.

The performances of two sampling schedules were evaluated. The rationale is that the design influences the bias and the accuracy of the population parameter estimates. The D-optimal design, which maximizes the determinant of the population Fisher information matrix, increases the accuracy and reduces the bias of the parameter estimates. The two different sampling schedules with similar amount of information achieved similar performances in the estimation of the population parameters in the face of BQL data. Comparison with the performances of a widely non-optimal design remains to be done.

REFERENCES

1. JP Hing et al. *J. Pharmacokin. Pharmacodyn*. 28 : 465-480 (2001)

2. SL Beal. *J. Pharmacokin. Pharmacodyn*. 28 : 48-504 (2001)

3. D. Spiegelhalter et al. *BUGS 0.5 Manual*. MRC Biostatistics Unit, Institute of Public Health, Cambridge, UK. 1996.

4. M. Tod et al. *J. Pharmacokin. Biopharm.* 26 : 689-716 (1998).

CONCLUSIONS

Conclusions for one- or two-compartment model are similar: Discarding BQL values leads to a poor estimation of the population parameters especially when the amount of information is low and when the proportion of BQL values is high.

Method 2 works quite well and is very easy to implement in many softwares, but its performances may depend on the imputed value, the standard deviation, and the terminal half-life.

Method 3 works only slightly better and is not easy to implement in most softwares (including NONMEM), unless WinBUGS.

Different sampling schedules with same information content yields similar performances.

TABLE I. CONDITIONS OF THE SIMULATIONS

Name	QL	N	m	n[a]	%BQL[b]	Schedule
Ia	**0.5**	**60**	**3/3**	**240**	**10**	T_D
Ib	0.5	60	3/3	240	10	T_{nonD}
IIa	**0.5**	**60**	**1/3**	**120**	**20**	T_D
IIb	0.5	60	1/3	120	20	T_{nonD}
IIIa	**0.5**	**180**	**1/3**	**360**	**20**	T_D
IIIb	0.5	180	1/3	360	20	T_{nonD}
IVa	**0.65**	**60**	**1/3**	**120**	**30**	T_D
IVb	0.65	60	1/3	120	30	T_{nonD}

QL = quantification limit. N = number of individuals. m = typical number of measures per individual among the first three samples.

[a] the total number of data points n includes the BQL data points

[b] percentage of BQL values with respect to n

FIGURE 1. PLOT OF A TYPICAL DATA SET FOR CASE IIA

The dashed line shows the quantification limit.

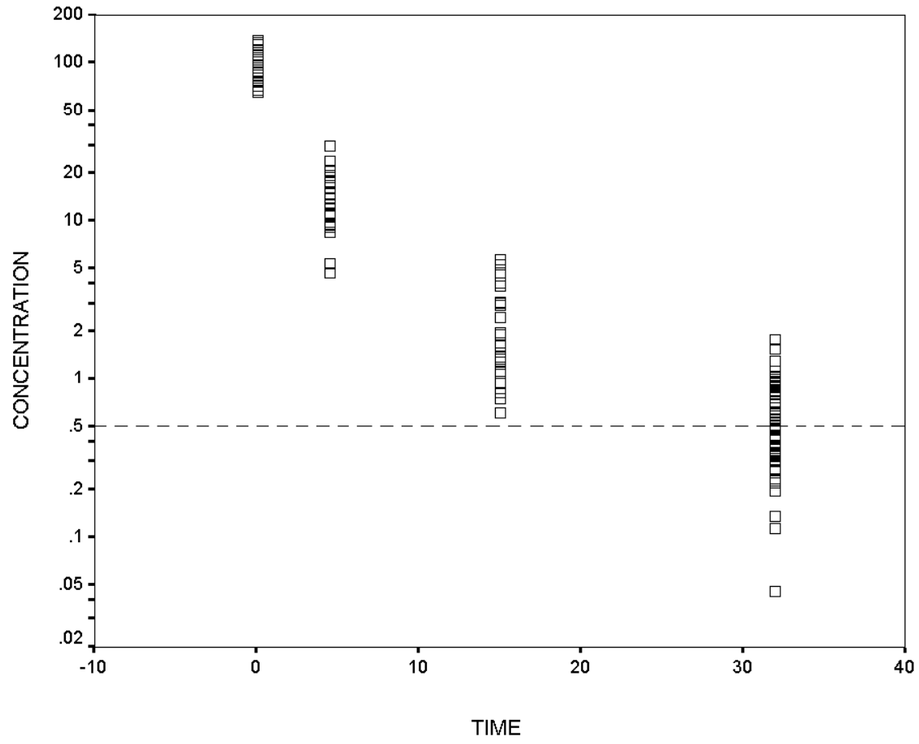
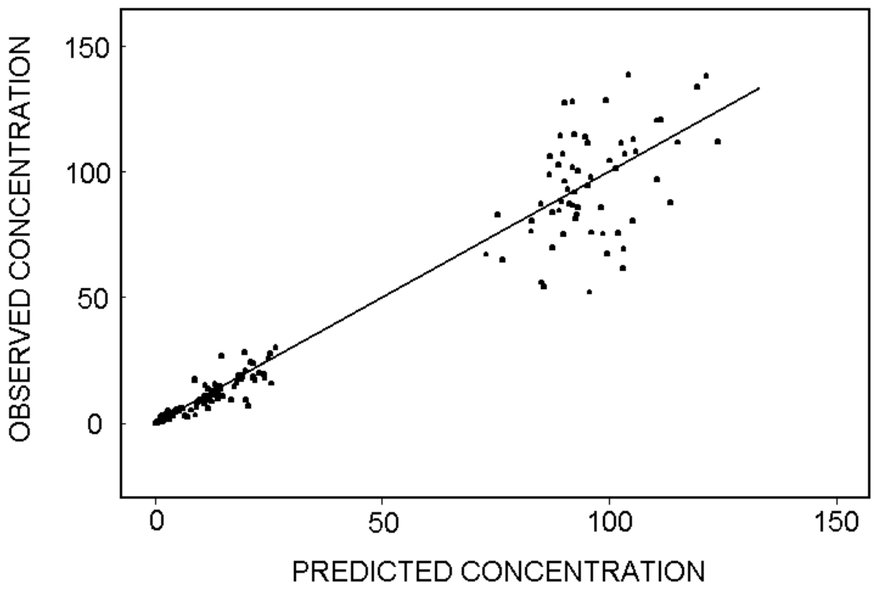
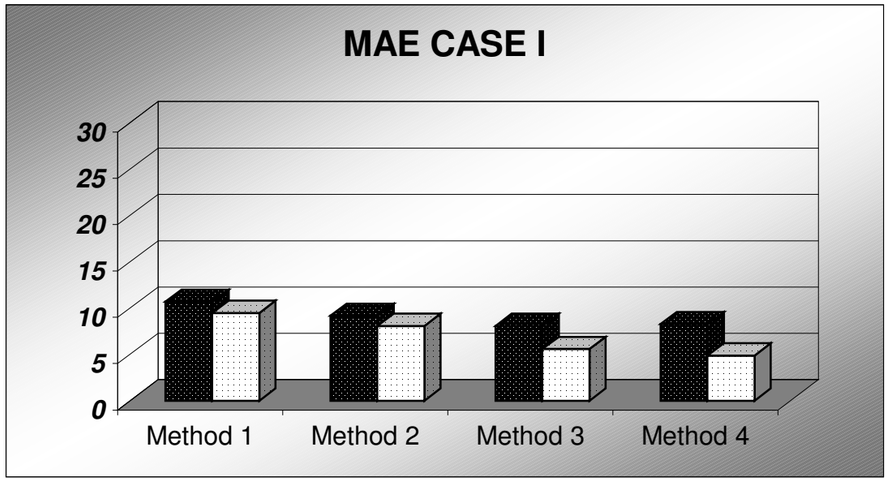
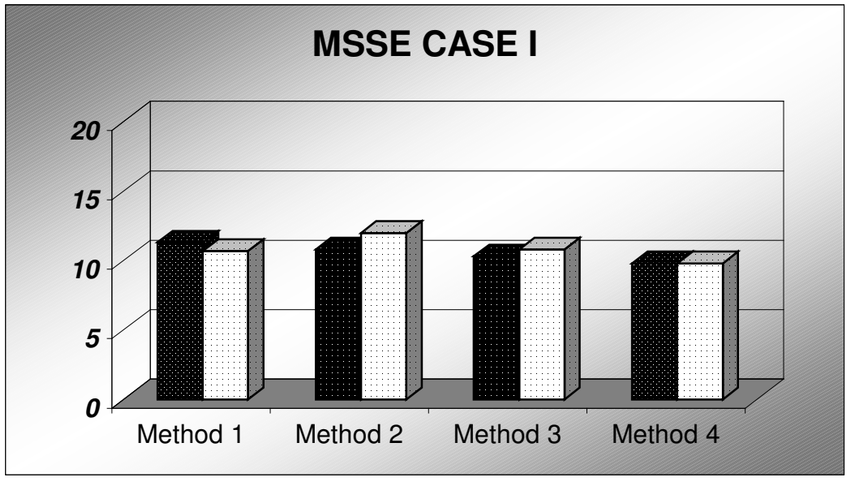
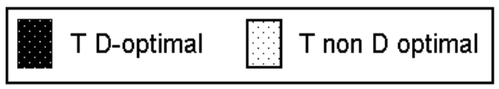


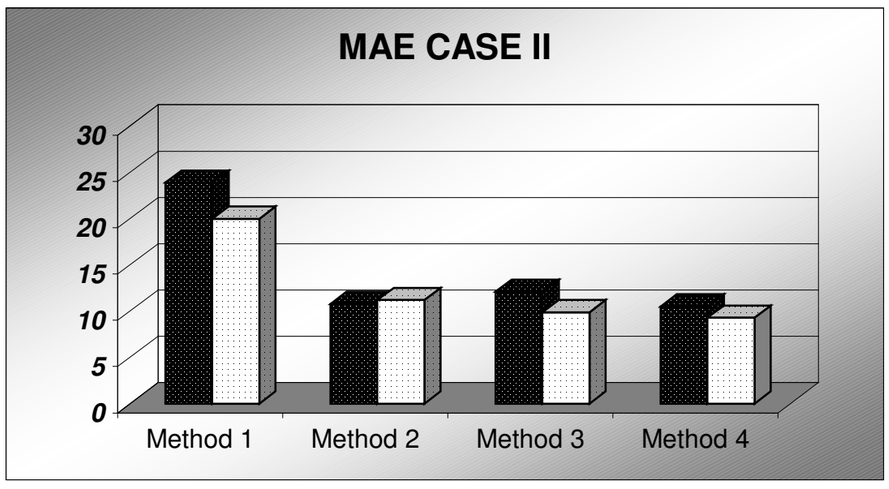
FIGURE 2. PLOT OF OBSERVED VERSUS PREDICTED CONCENTRATION FOR A RUN OF CASE IA

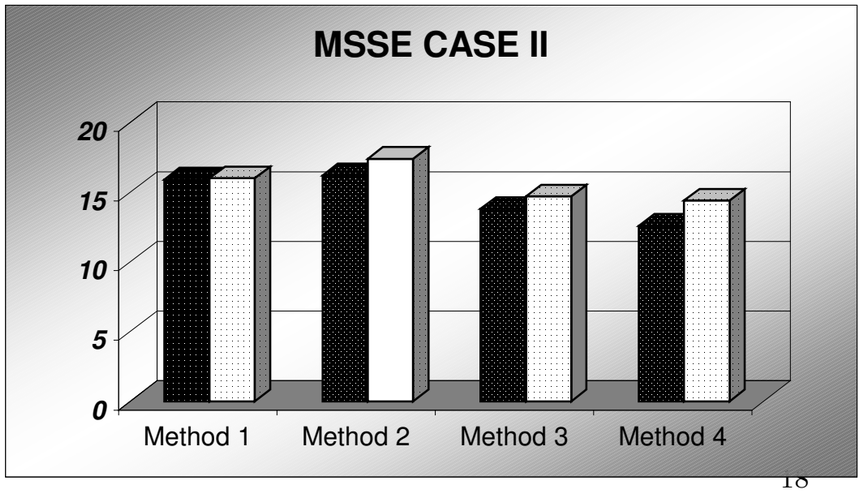


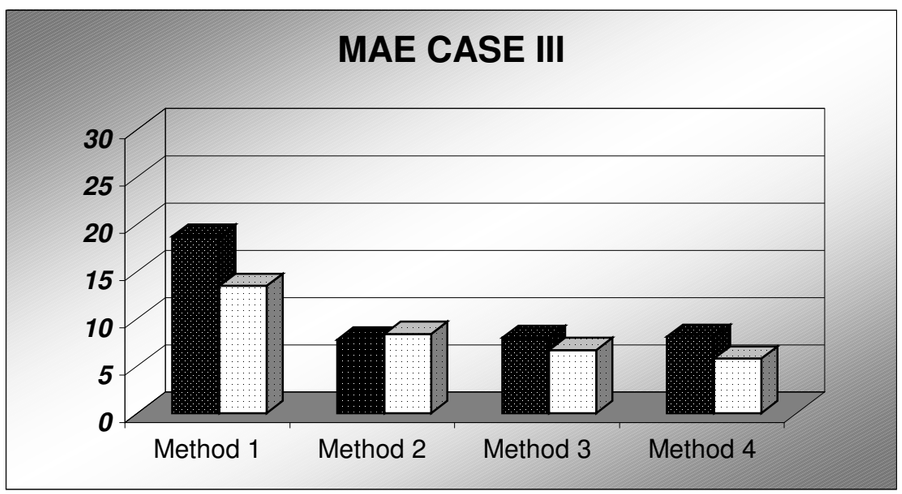

MAE CASE I
30
25
20
15
10
5
0
Method 1
Method 2
Method 3
Method 4

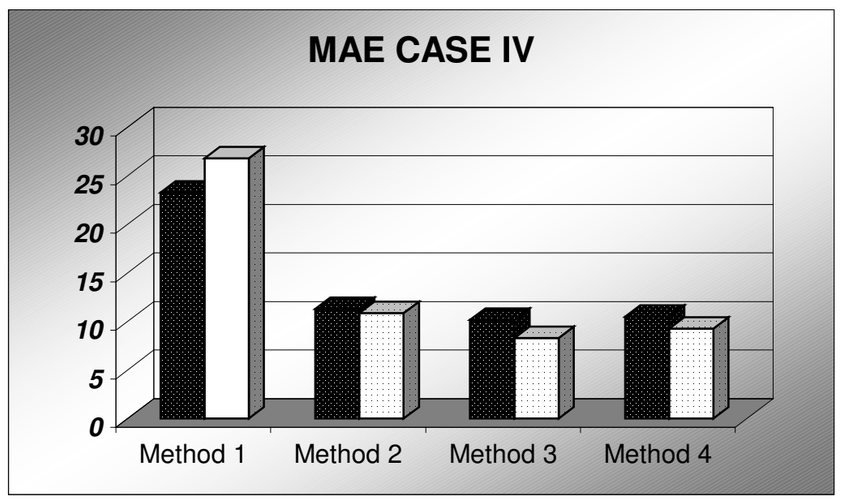

MSSE CASE I
20
15
10
5
0
Method 1
Method 2
Method 3
Method 4

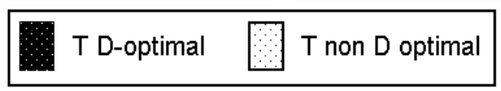

T D-optimal
T non D optimal

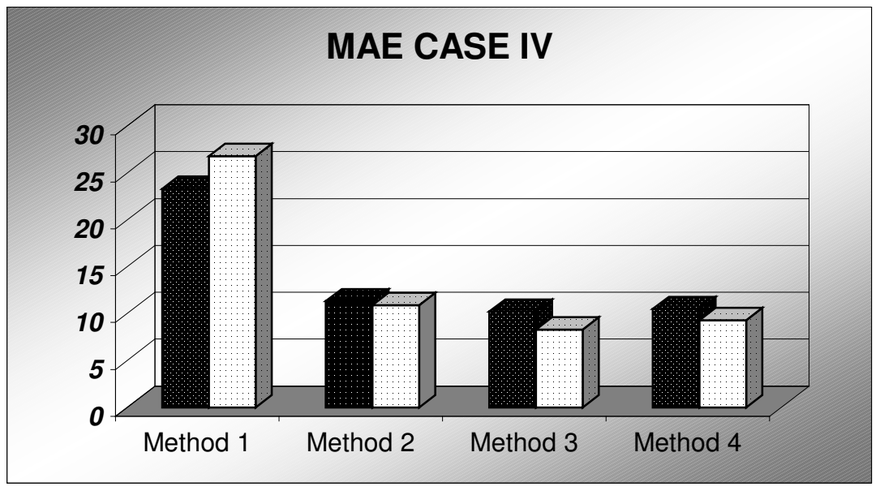

MAE CASE II
30
25
20
15
10
5
0
Method 1
Method 2
Method 3
Method 4

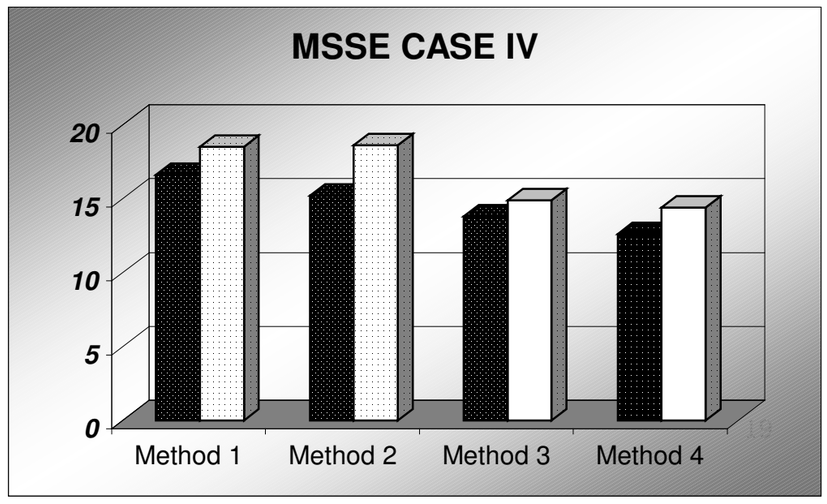

MSSE CASE II
20
15
10
5
0
Method 1
Method 2
Method 3
Method 4


MAE CASE III
30
25
20
15
10
5
0
Method 1
Method 2
Method 3
Method 4


MAE CASE IV
30
25
20
15
10
5
0
Method 1
Method 2
Method 3
Method 4


T D-optimal
T non D optimal


MAE CASE IV
30
25
20
15
10
5
0
Method 1
Method 2
Method 3
Method 4


MSSE CASE IV
20
15
10
5
0
Method 1
Method 2
Method 3
Method 4