DYNAMIC EVALUATION OF BATTERS IN NIPPON PROFESSIONAL BASEBALL LEAGUES

– A NEW PROPOSITION RELATING TO MULTIPLIER BOUNDS IN DEA-AR –

Tohru Ueda
Seikei University
3-3-1 Kichijoji-Kitamachi, Musashino, Tokyo 180-8633 Japan
ueda@st.seikei.ac.jp

Abstract: In each of Nippon Professional Baseball Leagues (NPBL), only about twenty batters over 220 times at bat played during years 2008-2012. Therefore in evaluation of batters, we cannot use long time series data for the same batter. At first possibility of efficiency evaluation based on short past data is discussed. In the next place evaluations of batters based on batting orders is discussed. In such DEA models as CCR, multipliers which may cause bad effects for evaluation may be neglected, that is, the values of the corresponding multipliers may become 0. To improve this shortcoming, the assurance region methods, which have bounds relating to multipliers were proposed. We propose new methods based on the absolute deviation, by which the bounds are derived effortlessly from limited information, i.e., partial ranking data. The methods are applied to the evaluation of baseball players in NPBL.

Keyword: DEA, assurance region, baseball, absolute deviation, super-efficiency

1. INTRODUCTION

In each of Nippon professional baseball leagues, only about twenty batters over 220 times at bat played through years 2008-2012. Therefore in evaluation of batters, we cannot use long time series data for the same batter and we use methods of forecasting based on Stein's estimator [1], [2] which shrinks forecasts towards the total mean. Accuracies of various forecasts are compared.

In evaluation of baseball batters, batting orders may have large influence. For example, for the first batter, hits may be more important than runs batted in, but for the fourth batter, runs batted in may be more important than hits. In Data Envelopment Analysis (DEA) many optimal weights (multipliers) for inputs and outputs may become zeros because the evaluated Decision Making Units (DMU) can obtain the efficiency score of 1 by neglecting inputs or outputs that are inferior to the inputs or outputs of other DMUs [5]. This means that if inputs or outputs showing the performances of DMUs are neglected, valuable information may consequently be lost.

To improve this shortcoming, the assurance region methods, which have bounds relating to weights were proposed ([3], [4], [6]-[9], [11]-[13]). However, reference [3] states, "No method is all-purpose and different approaches may be appropriate in different contexts" and reference [6] states, "There is no single correct process for determining numerical values of bounds".

We agree with these opinions and several researchers have proposed various methods of determining bounds. To determine bounds on weights, reference [6] discusses the use of regression analysis, references [12] and [13] discuss the use of canonical correlation analysis, and to set upper and lower bounds on weights in the "bounded" formulation, references [8] and [9] use weights which were obtained from unbounded runs of DEA. References [4] and [7] suggest the setting of bounds based on expert judgments, and reference [11] is a concrete of them. Reference [11] proposed a method that decides bounds by utilizing the judgments of people who know well the characteristics of the evaluated objects. Quantification of bounds is accomplished by Saaty's Analytic Hierarchy Process (AHP) [10] based on paired comparison results, but when the number of objects is M, $M(M-1)/2$ comparisons are needed and comparisons between unimportant objects are difficult in general. We can take fewer comparisons for rank order data than for paired comparison data and if ranking among unimportant objects can be avoided, ranking becomes easier.

We discuss cases where more important m ($<M$) objects than others are ranked and propose a method which does not use the ranking of all objects and transforms the ranking data into positive real numbers [14], where we used Maximization of the Variance Between Objects which has a possibility of local maximization. In this paper we use Maximization of the Deviation Between Objects which is formulated as a linear programming problem.

The proposed method is applied to the evaluation of batters in Nippon Professional Baseball Leagues.

2. HOW TO EVALUATE BATTERS

We evaluate baseball batters over 220 times at bat from the 2008 season to the 2012 season. We use the following items (M=9) as objects.

1: (runs batted in)/(plate appearances),
2: batting average, that is, hits/(at-bats),
3: (doubles+triples)/ (at-bats),
4: (stolen bases) /(plate appearances),
5: strikeouts/(plate appearances).
6: (sacrifice hits) /(plate appearances)
7: (sacrifice fries) /(plate appearances)
8: walks/(plate appearances)
9: (home runs) /(at-bats),

These items are used as outputs of DEA, where the best value and the worst value of each item are transformed into 1 and 0 in Sec.5, respectively. We must note that for item 5 the largest value is transformed into 0 and the smallest value is transformed into 1. Each DMU has single input and is set to 1. In the following, **item *i*** means **batting result *i***.

Basically for efficiency evaluation of baseball batters we use the CCR model. However since batting orders may have large influence in efficiency evaluation, for each batting order we use the assurance region methods, which have bounds relating to weights as follows.

(assurance region method)

$$\max\text{imize} \sum_{j=1}^{M=9} u_j y_{jo}$$

$$\text{subject to} \quad \sum_{j=1}^{M=9} u_j y_{jg} \le 1 \ \ (g = 1,\dots,n\,;\, g \ne o)$$

$$u_j L_{jk} \le u_k \le u_j U_{jk} \ \ (j \ne k)$$

$$u_j \ge 0 \ (j = 1,2,\dots,[M = 9])$$

where since single input is set to 1, $v\,x_g = v\,x_o = v = 1$ and introducing v_0 in virtual inputs is meaningless.

How to decide L_{jk} and U_{jk} is shown in Sec.5.

3. FORECASTING METHODS

Nippon professional baseball leagues have two leagues, Central League (Ce) and Pacific League (Pa). In each of them only about twenty batters over 220 times at bat played through years 2008-2012 (We got 23 batters by addition of batters who exceeded 220 times at bat in four seasons). Therefore in evaluation of batters, we cannot use long time series data for the same batter and we use methods of forecasting based on Stein's estimator [1], [2] which can be used even if there is only an observe for an parameter. Suppose that random variables x_i (i=1,…,p) are normally distributed: $N(\theta_i,\sigma^2)$. Stein's estimator is given by

$$\hat{\theta}_i = a\,x_i + (1-a)\,\bar{X}, \quad \bar{X} = \sum_{i=1}^{p} x_i \,. \qquad (1)$$

A parameter a is decided by an approximate of minimization of

$$\sum_{i=1}^{p} E(\hat{\theta}_i - \theta_i)^2 \,.$$

In this paper forecasts are calculated along the lines of Eq.(1). Let the result of batter i in year t be $x(t, i)$. An forecast of batter i in year (t+1) is given by

$$\hat{x}(t+1,i) = a\,x(t,i) + b\,x(t-1,i) + c\,\bar{X}(t) \quad (2)$$

or

$$\hat{x}(t+1,i) = a\,x(t,i) + c\,\bar{X}(t) \quad (3)$$

where $\bar{X}(t) = \sum_{i=1}^{p} x(t,i)/p$. Parameters a, b and c are decided by minimization of

$$\varepsilon_{t+1} = \sum_{i=1}^{p} |\,\hat{x}(t+1,i) - x(t+1,i)\,|, \quad p = 23 \cdot$$

Table 1 shows errors of forecasts in Pa-League where

10', 11', 12': use of observed values

10'F, 11'F, 12'F: use of forecasts obtained by Eq.(2) or (3)

$$abs(A-B) = \sum_{i=1}^{p} |\,A(i) - B(i)\,|$$

For example, abs(11'-11'F) is a sum of absolute forecasts errors for all batters of 11' and abs(11'-10') is a sum of absolute differences of observes for all batters between 11' and 10'. Here 11'F and 12'F use the same regression coefficients as obtained for 10'F.

If abs(11'-10')<abs(11'-11'F), forecasts cannot be used. Equation (3) is slightly more accurate than Eq.(2). Using forecasts obtained by Eq.3 and observes in year 2011 and 2012, efficiency scores were calculated. Results are shown in Table 2. The third and fifth columns are smaller than the second and fourth columns.

Table 2 : total absolute differences of efficiency scores

	abs(11'-10')	abs(11'-11F')	abs(12'-11')	abs(12'-12'F)
Ce	1.910	1.411	2.502	2.330
Pa	2.350	2.098	1.844	1.454

Table 1 : Errors of forecasts by each item

Pa_Eq.(2)									
item No.	1	2	3	4	5	6	7	8	9
abs(10'-10'F)	0.357	0.559	0.242	0.294	0.652	0.221	0.084	0.464	0.117
abs(10'-9')	0.416	0.759	0.374	0.345	0.720	0.262	0.109	0.466	0.165
abs(11'-11'F)	0.604	0.631	0.316	**0.349**	**0.653**	0.208	**0.088**	0.425	**0.222**
abs(11'-10')	0.651	0.751	0.381	**0.232**	**0.633**	0.220	**0.086**	0.441	**0.206**
abs(12'-12'F)	0.442	**0.630**	**0.336**	**0.274**	**0.449**	0.297	0.075	0.326	0.158
abs(12'-11')	0.470	**0.575**	**0.276**	**0.199**	**0.445**	0.329	0.095	0.331	0.167
a	0.592	0	0.168	0	0.410	0.369	0.031	0.946	0.623
b	0.210	0.466	0.288	0.946	0.590	0.631	0	0	0.208
c	0.197	0.534	0.544	0.054	0	0	0.969	0.054	0.168

Pa_Eq.(3)									
item No.	1	2	3	4	5	6	7	8	9
abs(10'-10'F)	0.375	0.670	0.259	0.334	0.720	0.260	0.084	0.464	0.127
abs(10'-9')	0.416	0.759	0.374	0.345	0.720	0.262	0.109	0.466	0.165
abs(11'-11'F)	**0.654**	0.658	0.315	**0.233**	0.633	**0.223**	**0.088**	0.425	**0.230**
abs(11'-10')	**0.651**	0.751	0.381	**0.232**	0.633	**0.220**	**0.086**	0.441	**0.206**
abs(12'-12'F)	0.422	0.557	**0.330**	0.179	0.445	0.322	0.075	0.326	0.154
abs(12'-11')	0.470	0.575	**0.276**	0.199	0.445	0.329	0.095	0.331	0.167
a	0.678	0.275	0.134	0.869	1	0.945	0.031	0.946	0.808
c	0.322	0.725	0.866	0.131	0	0.055	0.969	0.054	0.192

<3>

4. CCR EFFICIENCY SCORES IN ALL DATA

Using the CCR model, we obtained simultaneously efficiency scores of all baseball batters over 220 times at bat from the 2008 season to the 2012 season.

Figure 1 shows examples of efficiency scores for the same batters in Pa-League. This shows that changes of efficiency scores are large and this causes difficulties of forecasting. Therefore forecasting is not considered in the next section.

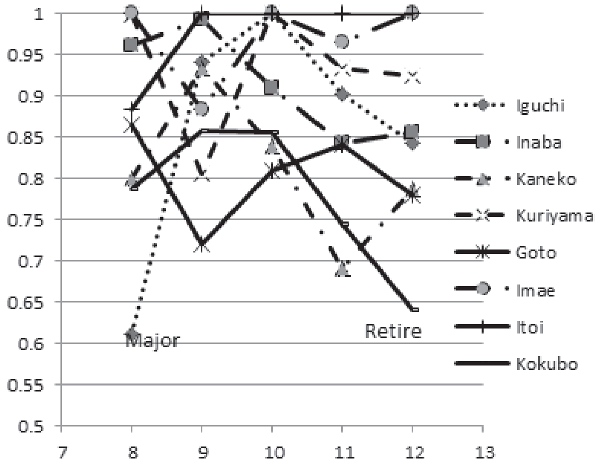

Figure 1: Examples of efficiency scores for the same batters

5. DERIVATION OF IMPORTANCE SCORES

5.1. Maximization of Between Variance

We would like to know importance of M objects (items) and we ask N persons to rank them. However, ordering all objects, especially ranking among unimportant objects, may be difficult. More important m ($<M$) objects than others are ranked. We give score t_1 for the most favorite object, t_2 for the second favorite object,…, t_m for the m-th favorite object, and t_{m+1} for non-selected objects. Variant rankings are usually obtained from person to person. Let a score of person $\boldsymbol{i}$ and object $\boldsymbol{j}$ be $\boldsymbol{f_{ij}}$. If person i answers object 2 as the most favorite object and object 5 as the second favorite object, object 2 is given score t_1, that is, $f_{i2}= t_1$, and object 5 is given score t_2, that is, $f_{i5}= t_2$ (See Table 3). Because we would like to know ratios among t_i, let t_{m+1}=1 and $\log_e t_k \equiv \ln t_k$ $\log_e t_i$ is discussed as it becomes familiar with mean and variance.

Table 3 : Ranking data and scoring image

Ranking data (a)

	rank1	rank2	rank3	rank4	rank5
p1	**2**	**5**	1	10	7
p2	2	1	5	10	3

Scoring image (b)

p1	**obj 1**	**obj 2**	**obj 3**	**obj 4**	**obj 5**
	$f_{11}=t_3$	$\boldsymbol{f_{12}=t_1}$	$f_{13}=t_6$	$f_{14}=t_6$	$\boldsymbol{f_{15}=t_2}$
	obj 6	**obj 7**	**obj 8**	**obj 9**	**obj 10**
	$f_{16}=t_6$	$f_{17}=t_5$	$f_{18}=t_6$	$f_{19}=t_6$	$f_{1,10}=t_4$
p2	**obj 1**	**obj 2**	**obj 3**	**obj 4**	**obj 5**
	$f_{21}=t_2$	$f_{22}=t_1$	$f_{23}=t_5$	$f_{24}=t_6$	$f_{25}=t_3$
	obj 6	**obj 7**	**obj 8**	**obj 9**	**obj 10**
	$f_{26}=t_6$	$f_{27}=t_6$	$f_{28}=t_6$	$f_{29}=t_6$	$f_{2,10}=t_4$

p : person, obj : object

In [14] differentiation among objects was realized through Maximization of the Variance Between Objects (Between Variance) under the constant Total Variance. that is, the following formulation, MV1:

$$\text{(MV1)} \quad \max\text{imize} \sum_{j=1}^{M} (\mu_j - \mu)^2$$

subject to

$$\sum_{j=1}^{M} \sum_{i=1}^{N} \{(\ln f_{ij}) - \mu\}^2 \,/\, (MN) = V_T \ : \text{constant}$$

$$\ln(t_h) - \ln(t_{h+1}) \geq C \geq 0, \ t_{m+1} = 1$$

where $\mu = \sum_{j=1}^{M} \sum_{i=1}^{N} (\ln f_{ij}) / (MN)\,;\ \mu_j = \sum_{i=1}^{N} (\ln f_{ij}) / N$

However, this solution may be the local optimum. To obtain the global optimum we will use the absolute deviation in the following subsection.

5.2. Maximization of Between Deviation

Suppose that Between Absolute Deviation, D_B, Within Absolute Deviation, D_W, and Total Absolute Deviation, D_T, are given as following:

<4>

Between Absolute Deviation: $D_B = \sum_{j=1}^{M} |\mu_j - \mu| / M$

Within Absolute Deviation:

$$D_W = \sum_{j=1}^{M} \{\sum_{i=1}^{N} |f_{ij} - \mu_j| / N\} / M$$

Total Absolute Deviation:

$$D_T = \sum_{i=1}^{N} \sum_{j=1}^{M} |f_{ij} - \mu| / (NM)$$

Relating to variance, Total Variance, V_T, is the sum of Between Variance, V_B, and Within Variance, V_W, but relating to absolute deviation, Total Absolute Deviation, D_T, is not usually equal to the sum of Between Absolute Deviation, D_B, and Within Absolute Deviation, D_W.

An absolute value of v, $|v|$, is treated by s_1+s_2, where

$$v - (s_1 - s_2) = 0,\ \ s_1 \ge 0, s_2 \ge 0$$

$$C_0(1-b) \ge s_1 \ge 0,\ \ C_0 b \ge s_2 \ge 0,\ \ b \in \{0,1\}$$

C_0 is a constant.

From the above mentioned facts, we can obtain the following linear programming formulation, AD1:

Objective： $\max \sum_{j=1}^{M} (s_{1j} + s_{2j}) / M$

Constraints： $\mu = \sum_{i=1}^{N} \sum_{j=1}^{M} f_{ij} / (NM), \mu_j = \sum_{i=1}^{N} f_{ij} / N$

$$\mu_j - \mu + s_{1j} - s_{2j} = 0; j = 1, 2, ..., M$$

$$\sum_{j=1}^{M} (s_{1j} + s_{2j}) / M - D_B = 0$$

$$M \cdot C_2 (1 - b_j) \ge s_{1j} \ge 0,\ \ M \cdot C_2 \cdot b_j \ge s_{2j} \ge 0, b_j \in \{0,1\}$$

$$f_{ij} - \mu_j + t_{ij}^{(1)} - t_{ij}^{(2)} = 0 \Leftarrow |f_{ij} - \mu_j| = t_{ij}^{(1)} + t_{ij}^{(2)}$$

$$\sum_{i=1}^{N} \sum_{j=1}^{M} (t_{ij}^{(1)} + t_{ij}^{(2)}) / (NM) - D_W = 0$$

$$t_{ij}^{(1)} \ge 0, t_{ij}^{(2)} \ge 0 : i = 1, ..., N; j = 1, ..., M$$

$$D_B + D_W = C_2$$

$$\ln(t_k) - \ln(t_{k+1}) \ge C_0 : k = 1, ..., m ; t_{m+1} = 1$$

Where N =10, M = 9, m = 6, C_0 = ln (9/20) and C_2 = DB+DW obtained when t_k= 10–k.

Table 4 shows evaluation of items (batting results) obtained from 10 persons. Using AD1, we obtained such values of t_k as Table 5. From t_k, we can calculate such the importance μ_j of item i in each batting order, j as Table 6. By Table 4 and Table 5, we can calculate such the ratio, r_{ab}, of scores on (item a vs. item b) as Table 8 where ratios among items with small μ_j are not considered. We use the minimum of r_{ab} as L_{ab}, and the maximum of r_{ab} as U_{ab} in the assurance rigion method. For example, in the first batter, $L_{2,4}$=0.896 and $U_{2,4}$=1.552.

Table 4 : Item number of rank *i* given by person (p) *h*

	rank 1	rank 2	rank 3	rank 4	rank 5	rank 6
p1	8	2	4	5	6	7
p2	2	4	5	8	6	3
p3	8	4	2	3	5	6
p4	8	2	5	4	6	3
p5	2	8	4	3	6	5
p6	2	5	8	4	6	3
p7	2	8	5	6	4	3
p8	5	2	8	3	4	6
p9	2	8	3	4	5	6
p10	8	2	4	6	5	3

Table 5 : Values of t_k in the first batter

t_1	t_2	t_3	t_4	t_5	t_6	t_7
8.687	7.783	6.973	6.248	5.598	5.015	1

Table 6 : μ_i of item i in each batting order, j

	batting order					
item	1	2	3	4	5	6,7,8
1	0	0.033	1.407	2.144	1.641	0.066
2	2.096	1.754	1.948	1.381	2.013	1.907
3	1.550	0.357	1.886	1.957	1.883	0.110
4	1.887	1.630	0	0	0	0.209
5	1.865	1.784	1.970	1.630	1.925	1.785
6	1.711	2.152	0.194	0	0.163	0.848
7	0.161	0.167	1.611	1.935	1.870	1.301
8	2.052	1.514	0.698	0.485	0.326	1.818
9	0	0	1.578	1.814	1.597	0

Table 7 : Scores of item *i* by P*h* in the first batter

	item 1	item 2	item 3	item 4	item 5	item 6	item 7	item 8	item 9
p1	1	7.783	1	6.973	6.248	5.598	5.015	8.687	1
p2	1	8.687	5.015	7.783	6.973	5.598	1	6.248	1
p3	1	6.973	6.248	7.783	5.598	5.015	1	8.687	1
p4	1	7.783	5.015	6.248	6.973	5.598	1	8.687	1
p5	1	8.687	6.248	6.973	5.015	5.598	1	7.783	1
p6	1	8.687	5.015	6.248	7.783	5.598	1	6.973	1
p7	1	8.687	5.015	5.598	6.973	6.248	1	7.783	1
p8	1	7.783	6.248	5.598	8.687	5.015	1	6.973	1
p9	1	8.687	6.973	6.248	5.598	5.015	1	7.783	1
p10	1	7.783	5.015	6.973	5.598	6.248	1	8.687	1

Table 8 : Values of (item 2 / item 4) in p*h* of Table 7

p1	p2	p3	p4	p5	p6	p7	p8	p9	p10
1.116	1.116	0.896	1.246	1.246	1.390	1.552	1.390	1.390	1.116

Table 9 : Top 20 fourth batters of efficiency scores in 2012

Ce

Abe	1.266
Sakamoto	0.935
Balentien	0.907
Milledge	0.855
Toritani	0.825
Blanco	0.819
N.Nakamura	0.794
Wada	0.793
Chono	0.738
Matsumoto	0.737
Soyogi	0.716
Eldred	0.716
Uemoto	0.712
Dobayashi	0.704
Murata	0.688
Kawabata	0.688
Ramirez	0.684
Morino	0.676
R.Arai	0.674
Amaya	0.671
Oshima	0.665

Pa

Matsuda	1.088
LeeDae-Ho	1.081
Masuda	1.025
Nakajima	1.012
Whitesell	1.008
Pena	0.987
Imae	0.976
T.Nakamura	0.950
Uchikawa	0.932
Kadonaka	0.906
Osaki	0.868
Inaba	0.848
Nakata	0.843
Makita	0.836
Kataoka	0.826
Baldiris	0.824
Kiyota	0.821
Ginji	0.811
Iguchi	0.805
German	0.793
Asamura	0.778

Table 10 : Top 20 third batters of efficiency scores in 2012

Ce

Abe	1.428
Sakamoto	0.966
Wada	0.917
Toritani	0.837
Milledge	0.830
Chono	0.818
Ishihara	0.809
Ramirez	0.808
Balentien	0.807
N.Nakamura	0.804
Fujimura	0.804
Uemoto	0.797
Tanaka	0.795
Oshima	0.791
Kawabata	0.786
Yamato	0.765
Araki	0.764
Soyogi	0.762
Matsumoto	0.760
Amaya	0.756

Pa

Matsuda	1.160
Uchikawa	1.088
Nakajima	1.083
Imae	1.066
Masuda	1.065
LeeDae-Ho	1.048
Itoi	1.012
Ginji	1.003
Kadonaka	1.003
Osaki	0.995
Tanaka	0.959
Nemoto	0.947
Kuriyama	0.925
Kiyota	0.912
Whitesell	0.911
Baldiris	0.908
Inaba	0.902
Pena	0.892
German	0.884
Akiyama	0.883

6. DERIVATION OF AR EFFICIENCY SCORES

We can calculate efficiency scores by the assurance region method (AR) described in Sec.2, where by addition of $g \neq o$, efficiency scores can become larger than 1. We obtained such efficiency scores of each batting order in each year as Table 9.

7. ASSIHNMENT OF POSITIONS

We select the best eight where not only batting result but also positions are considered as follows:

(assignment problem of positions)

$$\max \sum_{i,k,j} I(i,k,j) \times S(i,k)$$

subject to

$$\sum_{i,k} I(i,k,j) \times A(i,j) = c_j ; c_1 = 1, c_2 = 1, c_3 = 3, c_4 = 3$$

$$\sum_{i,j} I(i,k,j) \times A(i,j) = 1 : k \neq 6$$

$$\sum_{i,j} I(i,6,j) \times A(i,j) = 3, \quad \sum_{i,k,j} I(i,k,j) = 8$$

$$\sum_{k,j} I(i,k,j) \leq 1, \quad I(i,k,j) \in \{0,1\}$$

where $I(i, k, j)$ =1 : when batter i is selected as the k-th batter with position j

= 0 : otherwise

Position 1: catcher, 2: first baseman,

3: infielder except for the first baseman,

4: outfielder

The sixth batter is one of the sixth, seventh or eighth.

$A(i, j) = 1$: when batter i can take position j

= 0 : otherwise

$S(i, k)$: an efficiency score of batter i as the k-th batter which is obtained in Sec. 6

In this formulation the best eight in the Central League of 2012 was as follows.

<6>

Table 11 : Top 20 fourth batters of efficiency scores in 2011

Ce

Blanco	1.283
Kurihara	0.928
Brazell	0.926
Sledge	0.903
T.Arai	0.888
Balentien	0.878
Hatakeyama	0.874
Abe	0.873
Toritani	0.832
Ramirez	0.811
Murata	0.796
Chono	0.790
Wada	0.775
Y.Takahashi	0.762
Harper	0.745
Whitesell	0.740
Kawabata	0.737
Hirose	0.733
Kinjo	0.705
Hirata	0.697

Pa

T.Nakamura	1.301
Uchikawa	1.088
Nakajima	0.980
Matsunaka	0.913
Nakata	0.877
Matsuda	0.843
Imae	0.808
Iguchi	0.808
Fernandes	0.805
Itoi	0.793
Kuriyama	0.788
Kiyots	0.762
Baldiris	0.759
T.Okada	0.744
Kokubo	0.709
LeeDae-Ho	0.703
M.Nakamura	0.696
Garcia	0.691
Goto	0.690
Honda	0.685

(Initial result of the best eight in the Central League of 2012)

The first batter (#1) : Akamatsu as an outfielder

The second batter (#2): Fujimura as an infielder

The third batter (#3): Wada as a first baseman

The fourth batter (#4): Sakamoto as an infielder

The fifth batter (#5): Abe as a catcher

The sixth batter (#6): Tanaka as an infielder

The sixth batter (#6): Matsumoto as an outfielder

The sixth batter (#6): Uchimura as an outfielder

The fourth batter and fifth batter should be exchanged in my opinion. In the following, the fourth batter is selected in advance as one with the highest efficiency score and other batting orders are decided according to the above mentioned formulation.

(The best eight in the Central League of 2012)

The first batter (#1): Akamatsu as an outfielder

Table 12 : Top 20 third batters of efficiency scores in 2011

Ce

Chono	1.043
Miyamoto	1.004
Toritani	1.191
Kurihara	1.014
Blanco	1.314
Abe	1.005
Hirose	1.039
Brazell	0.960
Murton	0.950
Kinjo	0.949
Aoki	0.937
Kawabata	0.933
Ramirez	0.928
Hatakeyama	0.913
T.Arai	0.912
Hirano	0.898
Tanaka	0.893
Wada	0.882
Hirata	0.879
Sledge	0.841

Pa

Uchikawa	1.186
T.Nakamura	1.128
Itoi	1.045
Matsunaka	1.035
Honda	1.012
Nakajima	0.991
Kuriyama	0.981
Matsuda	0.937
Hara	0.923
Iguchi	0.915
Imae	0.898
Taguchi	0.861
Hasegawa	0.861
Baldiris	0.856
Asamura	0.848
Tanaka	0.846
M.Nakamura	0.838
Sakaguchi	0.836
Kiyota	0.830
Uchimura	0.824

The second batter (#2): Fujimura as an infielder

The third batter (#3): Wada as a first baseman

The fourth batter (#4): Abe as a catcher

The fifth batter (#5): Sakamoto as an infielder

The sixth batter (#6): Tanaka as an infielder

The sixth batter (#6): Matsumoto as an outfielder

The sixth batter (#6): Uchimura as an outfielder

(The best eight in the Pacific League of 2012)

The first batter (#1): German as an outfielder

The second batter (#2): Nemoto as an infielder

The third batter (#3): Uchikawa as an outfielder

The fourth batter (#4): Matuda as an infielder

The fifth batter (#5): Lee Dae-Ho as a first baseman

The sixth batter (#6): Nakajima as an infielder

The sixth batter (#6): Shima as a catcher

The sixth batter (#6): Masuda as an outfielder

(The best eight in the Central League of 2011)

<7>

The first batter (#1): Toritani as an infielder
The second batter (#2): Fujimura as an infielder
The third batter (#3): Abe as a catcher
The fourth batter (#4): Blanco as a first baseman
The fifth batter (#5): Kurihara as an infielder
The sixth batter (#6): Chono as an outfielder
The sixth batter (#6): Kimura as an outfielder
The sixth batter (#6): Hirose as an outfielder

(The best eight in the Pacific League of 2011)
The first batter (#1): Itoi as an outfielder
The second batter (#2): Honda as an infielder
The third batter (#3): Matsunaka as a first baseman
The fourth batter (#4): T. Nakamura as an infielder
The fifth batter (#5): Uchikawa as an outfielder
The sixth batter (#6): Nakajima as an infielder
The sixth batter (#6): Kuriyama as an outfielder
The sixth batter (#6): Satozaki as a catcer

8. CONCLUSION

Year by year batting results are varying. Therefore even if forecasts values are used, forecasting errors cannot be altered sharply.

The best eight were selected satisfactorily by prior decision of the fourth batter. The similar results may be obtained by modification of the objective function.

9. REFERENCES

[1] Stein C, Inadmissibility of the usual estimator for the mean of a multivariate distribution, *Proc. Third Berkeley Symp. Math. Statist. Prob.*, **1**, pp. 197–206 (1956).

[2] James W and Stein C, "Estimation with quadratic loss", *Proc. Fourth Berkeley Symp. Math. Statist. Prob.*, **1**, pp. 361–379 (1961).

[3] Allen R, et al., Weights restrictions and value judgments in Data Envelopment Analysis: Evolution, development and future directions, *Annals of Operations Research*, 73, 13-34 (1997).

[4] Beasley J, Comparing university departments. *Omega*, 18(2), 171-183 (1990).

[5] Cooper WW, Seiford LM and Tone K, *Data Envelopment Analysis*, 2nd Ed., Springer (2007).

[6] Dyson RG and Thanassoulis E, Reducing weight flexibility in data envelopment analysis. *Journal of the Operational Research Society*, 39(6), 563-576 (1988).

[7] Kornbluth JSH, Analyzing policy effectiveness using cone restricted Data Envelopment Analysis. *Journal of the Operational Research Society*, 42(12), 1097-1104 (1991).

[8] Roll Y, Cook WD and Golany B, Controlling factor weights in Data Envbelopment Analysis, *IIE Transactions*, 23(1), 2-9 (1991).

[9] Roll Y and Golany B, Alternate methods of treating factor weights in DEA, *Omega*, 21(1), 99-109 (1993)

[10] Saaty TL Analytic Hierarchy Process, McGraw-Hill (1980).

[11] Takamura Y and Tone K, A comparative site evaluation study for relocating Japanese government agencies out of Tokyo, *Socio-Economic Planning Sciences*, 37, 85-102 (2003).

[12] Ueda T, Objective restriction on multipliers in Data Envelopment Analysis (DEA), *Technical Reports of Seikei University*, 37(1), 27-31 (2000) (in Japanese)

[13] Ueda T: Application of multivariate analysis for DEA. Proceedings of DEA Symposium 2007, 96-101 (2007).

[14] Ueda T and Amatatsu H : Determination of Bounds in DEA Assurance Region Method- Its Application to Evaluation of Baseball Players and chemical companies. The journal of operations research of Japan, 52(3), 453-467 (2009).