韓國 軍事運營分析 學會誌
第15巻 第1號, 1989. 6. 30.

Stochastic Analysis to Characterize A CARMONETTE Data

Jae-Yeong Lee*
James C. Kolding**

Abstract

Events that occur within a high "resolution" combat model often need to be characterized and structured for representation in other models or for detailed analysis purposes. This paper attempts to characterize one of these events, helicopter deaths. The data analyzed for this paper were generated by a high resolution production simulation system, CARMONETTE. The thesis objective is to develop a model to characterize the event of interest, and check the fit of the developed model using a second set of data. The exponential model developed provides not only excellent characterization of Blue helicopter attrition but also sufficient confidence in our results for the purpose of aggregated combat simulation.

1. Introduction

A. CARMONETTE Model

1) purpose

CARMONETTE is Computer Simulation of Small Unit Combat Model which is used to analyze battalion-level combat doctrine and tactics.

* System Analysis Dept. in Army HQs
** U.S. Army Major in TRAC-White Sands.

2) Description

CARMONETTE is a computerized, stochastic, event-sequenced simulation of a combined arms air or ground war game. It is played on a variable terrain representation of grid squares at 100 meters resolution for an hour of combat engagement. Force representation of infantrymen or various vehicles including tanks, armored personnel carriers, air defense, and helicopters is at the individual up to platoon group size in a battalion-level force. Events pertaining to surveillance consider the effects of battlefield obscuration including weather, aerosol smoke, and artillery dust. Probabilities of hit and kill consider the biased dispersion of weapon systems based on moving firer/targets. Output consists of displays and detailed reports including the killer/victim scoreboard. (Ref. 1 : p. c-33)

This system is used by the US Army Training and Doctrine Command (TRADOC) at White Sands, New Mexico. CARMONETTE has been validated and verified for its ability to accurately portray combat. It typically simulates a 20 to 30 minute battle using a bottom-up approach, i. e. it simulates the individual weapon systems, their movements, observations and actions, and aggregates them.

CARMONETTE uses a Monte Carlo process for nearly all event generation. For any small time interval there are many thousand of Monte Carlo draws required to generate a single action or event on the battlefield. Due to these facts the investigation of a special event class, helicopter deaths, will not reflect individual embedded parameters of CARMONETTE.

B. Data Set

The data used for analysis were compiled from 20 separate simulation runs each using a different seed for the generation of the random numbers. The analysis of this paper used one half of the data for the model development phase and the other half of the data for the model verification via cross validation.

The original data obtained from White Sands is shown in Figure-1. Five rows of the data are listed and are representative of the 4838 rows of data used. The nineteen columns each contain informations that describe the events of the helicopter battle. Only five columns were needed to fully describe and characterize the deaths of the Blue helicopters. A legend for these five colums is listed in Figure-2 which is the cleaned data used in the model development. The second half of the data shown in Figur-3 was similarly cleaned for use in the cross validation process.

C. Scope of Study

The study objective is to model the loss of Blue (defenders) helicopters in this specific set of CARMONETTE simulation. To approach this goal, this study is confined to obtain three requirements. Requirement one is to clean the available data in order to address the question at hand. Since the CARMONETTE history file contained much more information than needed 90% of the original data was not used. The second requirement is to develop a model to characterize the event of interest. Thirdly, check the fit of the developed model using a second set of data.

21	7.7168	32	49	7.7168	2	182	14500	9100	1
1	0.59	0	0	0	282.1	0	1	3 ···	
21	7.7168	32	49	7.7168	2	182	14500	9100	2
67	0.099	109	8875	10090	280	5712	0	0 ···	
21	7.7168	32	49	8.4094	2	182	14500	9100	3
67	0.168	109	9525	9946	279.7	5046	0	0 ···	
21	7.7168	32	49	8.5779	2	182	14500	9100	4
67	0	109	9684	9912	279.6	4884	2.8	100000 ···	
21	7.7168	32	49	99.9999	0	0	0	0	0
0	0	0	0	0	0	0	0	0 ···	

Figure-1 Original Data Set

(4838 by 19 Matrix : five sample rows only are showed)

21	22.6826	2	108	200000
22	8.4517	2	103	200000
23	19.4529	2	107	200000
23	20.1824	2	106	200000
23	24.2051	2	108	200000
24	8.4536	2	109	200000
24	23.7117	2	108	200000
24	24.5557	2	106	200000
25	8.2974	2	109	200000
25	13.0762	2	101	200000
25	21.8748	2	108	200000
28	13.0339	2	101	200000
28	19.1589	2	107	200000
28	23.0969	2	108	200000
30	19.9436	2	106	200000
30	21.0972	2	107	200000
30	22.7410	2	108	200000
30	24.4514	2	101	200000

Figure-2 Model Development Data Set : Data Set 1
(Blue helicopter kills only : 18 deaths)

Column 1 : Replication number (replication 21 thru 30 were used)
Column 2 : Time from start of simulation (t=0) till event occurs
Column 3 : Player making the shot ; 1-Blue, 2-Red
Column 4 : Player number who is receiving fire, Blue helicopters are 101, 102, 103, 105, 106, 107, 108, 109
Column 5 : Effect of the shot ; 200000=killed

31	21.7332	2	108	200000
31	24.0342	2	107	200000
32	24.4919	2	108	200000
33	13.0247	2	101	200000
33	22.7791	2	107	200000
33	22.5784	2	105	200000
35	8.0476	2	109	200000
36	8.1267	2	109	200000
37	8.2979	2	109	200000
37	21.0327	2	108	200000
37	20.8997	2	106	200000
37	22.7336	2	107	200000
38	16.4292	2	108	200000
38	23.8584	2	107	200000
39	22.7910	2	108	200000
39	24.8953	2	101	200000
40	8.0750	2	109	200000
40	21.0371	2	107	200000
40	22.8516	2	108	200000
40	24.5508	2	106	200000

Figure-3 Model Fit Data Set : Data set 2
(Blue helicopter kills only : 20 deaths)

Column 1 : Replication number (replication 31 thru 40 were used)
Column 2 : Time from start of simulation (t=0) till event occurs
Column 3 : Player making the shot ; 1-Blue, 2-Red
Column 4 : Player number who is receiving fire, Blue helicopters are
101, 102, 103, 105, 106, 107, 108, 109
Column 5 : Effect of the shot ; 200000=killed

2. Model Development

A. Battle Scenario

The scenario played in CARMONETTE is one of NATO forces defending in prepared defensive positions against an attack by Warsaw Pact forces. The Blue force, defender, consists of a reinforced battalion. The principle reinforcement, as far as this analysis is concerned, is one company of attack helicopters, i. e. 5 attack systems and 3 scout/observation systems (8 systems in total). The attacking Red force is an armored regiment augmented by 8 attack helicopters. The simulation is terminate after minutes of battle time.

Throughout the conduct of the battle, CARMONETTE logs selected events, times and conditions in a post processor history file. The history file is the source of the data used in this paper.

B. Model Development Procedure

The first step was to understand the real world events the simulation was trying to portray. Since study goal is conerned with the death times of the Blue helicopters while in battle a briefreview of the detailed scenario and history file showed that the battle does not start at the beginning of the simulation. The simulation starts with the Blue forces starting their prepared defensive positions and the Red forces in their attack from their starting positions. It takes a random amount of time for the combatants to engage. This is also true for the Blue helicopters under investigation.

Further review of the history file showed, that the first appearance of Blue helicopter was always being shot at by a Red weapon system. So, in the analysis, this event signals the start of the battle for the Blue helicopters. Figure-4 shows the helicopter battle start times for both sets of data. Based on Figure-2 and 4, a replication case (25th replication) was drawn to show which ones are censored (i. e., survived) in Figure—5. Three helicopters out of eight are uncensored (i. e., killed) in a given time 25 minutes, and battle start time is 7. 6543 minutes.

To characterize the given data two ways of approach were considered; Poisson process and exponential model. Following two sections are dealing with those topics respectively.

Replication	Start Time
21	7.7168
22	7.7195
23	7.7434
24	7.6963
25	7.6543
26	7.7178
27	7.7424
28	7.6709
29	8.0227
30	7.7542

Start Times of Model Fit Data

Replicatioı	Start Time
31	9.1206
32	7.6614
33	7.7710
34	9.1130
35	7.6670
36	7.7129
37	7.6531
38	7.6477
39	8.0088
40	7.7024

Figure-4 Start Times of Model Development Data

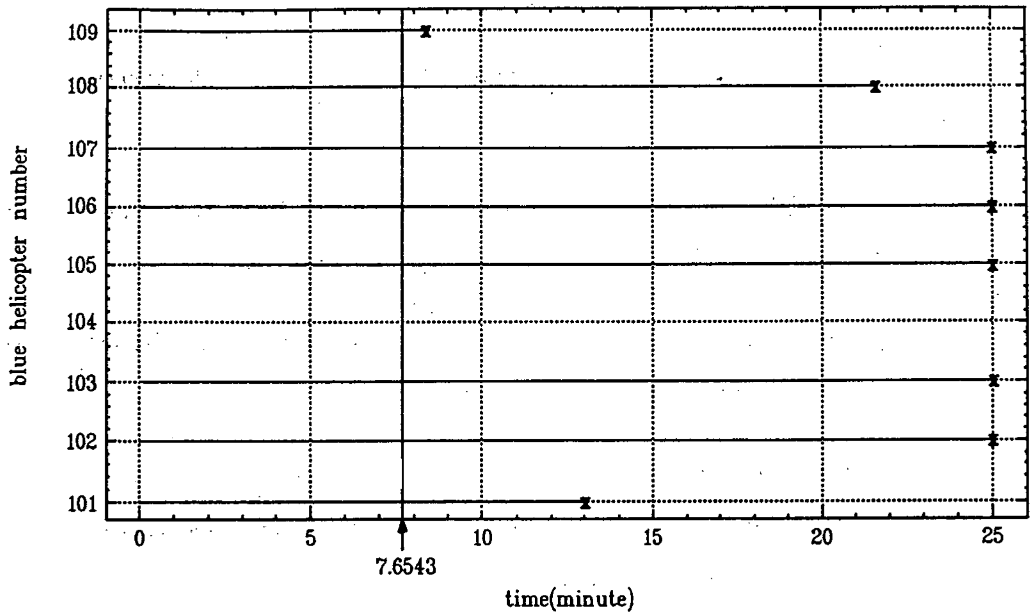

Figure-5 Time-Censored Data (25th replication case : d=3)

C. Poisson Process

The first idea was to model the helicopter deaths as Poisson process of the form, i.e. helicopter deaths occur according to a Poisson process with rate λ this could be done based on the following assumptions. (Ref. 3 : p. 197)

(i) if $N(t)$ represents the total numer of "Blue helicopter killed" that have occurred up to time t, then $N(0)=0$

(ii) The process has independent increments

(iii) The number of events in any interval of length t is

Poisson distributed with mean λt, that is, for all $s, t > 0$

$$P(N(t+s) - N(s) = n) = e^{-\lambda t}\frac{(\lambda t)^n}{n!}, \quad n=0, 1, 2, \cdots$$

But a closer look at the data revealed, that

- there are only eight helicopters in the battle which result in a quite abnormal Poisson process
- out of these eight helicopters only a minor fraction get killed (see Table-1), which reduces the amount of data to actually compute the rate, and
- the information about the survivors would be hard to incorporate into the model, which would reduce the significance of the results.

Due to these observations the idea of a Poisson process was not pursued. And the Poisson process, if actually valid, would not have provided any insight deeper than that the helicopters are taking chances out there.

Table-1 Number of Deaths and Survivors In Each Replication

Rep. No.	Uncensored	Censored	Rep. No.	Uncensored	Censored
21	1	7	31	2	6
22	1	7	32	1	7
23	3	5	33	3	5
24	3	5	34	0	8
25	3	5	35	1	7
26	0	8	36	1	7
27	0	8	37	4	4
28	3	5	38	2	6
29	0	8	39	2	6
30	4	4	40	4	4

D. Exponential Model

1) General

Assume an exponentially distributed lifetime for the individual helicopter. The survival function of the individual helicopter then is

$$\overline{F}(t) = e^{-\lambda t} \qquad (1)$$

with the "momoryless property". (Ref. 3 : p. 190)

This assumtion is based on the following ideas :

The helicopters are only a fraction of the friendly (Blue) forces. So the situation, as seen from the enemy's view, is a target-rich environment no matter how many helicopters have been killed. Or, starting it a diferent way, the enemy's concern has to be to fight the whole battle opposed just to hunting the helicopters. In the latter case, the assumption of identical survival functions for all helicopters would not hold. Helicopters fight in cycles of engagement and coverage : this can be described as entering and exiting the battle many times. Each entrance is independent of the previous actions. So it does not make a difference whether the helicopter enters for the 1st or for the i-th time. The assumed constant hazard rate is a weighted average of a low hazard rate in the phases of coverage and higher one during the times of exposure. The weights itself are dictated by what might be called battle intensity, which over the approximately 25 minutes of battle duration is assumed to be constant.

The model developed using the above assumption is one of eight items, Blue helicopters, on test during ten independent replication. since each replication is independent, the ten replication can be aggregated. And since the exponential model is used it dose not make a difference if we consider eight items on test for ten times or eighty items on test for one time.

All replications end at the same time (25.0 minutes) but not all helicopter battle start at the same time, as indicated in Figure-4. The helicopter battle start times are quite close, however which leaves us with a choice to either use the actual times on test or just take the minimun start time in the ten replications as common start time and procede from there.

Another option would be to average the start times over the ten replications, but then some negative life times might occur. In either case a situation of "time-censored" is predicted and the Maximun Likelihood Estimator (MLE) is exactly the same until it comes to substituting in the actual numbers. The first approach should produce the better result.

2) Failure—Censored and Time—Censored Cases

Note that life-testing experiments are usually destructive in that the items are destroyed at the end of the experiment and can not be used again. This limits the number of items we can test. we may put n items on test and terminate the experiment when a pre-assigned number of items, say $r(<n)$ have failed. The sample obtained from such an experiment are called "failure-censored" samples. Failure-censored sampling is almost mandatory in dealing with high cost sophisticated items such as colour television tubes. Another factor that affects the life—testing experiment is the amount of time required to obtain the complete sample. To limit this factor, we may put n items to test and terminate the experiment at a pre-assigned time t_0. The samples obtained from such an experiment are called "time-censored" samples. Time-censored sampling is almost essential in dealing with life-testing experiments in which the cost of experiments increases heavily with time.

In the failure-censored case data consist of life times of the r items that failed (say $x_{(1)}<x_{(2)}< \cdots < x_{(r)}$) and the fact that (n-r) items have survived beyond $x_{(r)}$. In the time-censored case data consist of the life times of items that failed before the time t_0 say $x_{(1)}<x_{(2)}< \cdots < x_{(m)}$ assuming that m items failed before t and tha fact that (n-m) items have survived beyond t_0. In the time-censored case t_0, the time of termination, is fixed while m, the number of items that failed before t_0, is a random variable. In the failure-censored case the situation is reverse in that r, the number of items that failed, is fixed, while $x_{(r)}$, the time at which the experiment is terminated, is a random variable. (Ref. 2 : p. 18-19)

3) Maximum Likelihood Estimator

This section shows the procedure to get M. L. E. for the hazard rate.

Let t_i=time from starting the battle until the i-th Blue helicopter killed (for uncensored data).

Let t_j=time from starting the battle until censoring time (for censored data).

The Blue helicopter's life times, t_i' s, are independent and identically distributed with the above stated survival function (equation 1) which translates into the distribution function F;

$$F(t, \lambda) = \begin{cases} 1-\exp(-\lambda t), & t \geqq 0 \\ 0, & \text{otherwise} \end{cases} \tag{2}$$

and the density function f;

$$f(t, \lambda) = \lambda \exp(-\lambda t) \qquad t \geqq 0. \tag{3}$$

Let $\underbrace{t_1,\ t_2,\ \cdots,\ t_d,}_{\text{uncensored (d)}}\ \underbrace{t_{d+1},\ \cdots,\ t_n}_{\text{censored (n-d)}}$

be the observed lifetimes of the helicopters. Then the M. L. E. for the t_i is:

$$\begin{aligned} L(\lambda) &= \Pi_{i=1}^{d} f(t_i, \lambda) \cdot \Pi_{j=d+1}^{n} (1-F(t_j, \lambda)) \\ &= \Pi_{i=1}^{d} \lambda \exp(-\lambda t_i) \cdot \Pi_{j=d+1}^{n} \exp(-\lambda t_j) \end{aligned} \tag{4}$$

Taking the natural logarithm and differentiate ln L(λ) with respect to λ we get

$$\ln L(\lambda) = d \ln \lambda - \lambda \sum_{i=1}^{d} t_i - \lambda \sum_{j=d+1}^{n} t_j \tag{5}$$

and

$$\hat{\lambda} = \frac{d}{\sum_{i=1}^{d} t_i + \sum_{j=d+1}^{n} t_j} \tag{6}$$

Then the mean survival time for Blue helicopters is

$$\hat{\mu} = \frac{1}{\hat{\lambda}} . \tag{7}$$

To approximate confidene intervals for the estimated parameter $\hat{\lambda}$, the Fisher information $i(\hat{\lambda})$ is used (Ref. 4)

$$i(\hat{\lambda}) = -\frac{d^2}{d\lambda^2}\ln(\lambda) = \frac{d}{\hat{\lambda}^2} . \tag{8}$$

The lower 95% confidence bound for the parameter is approxmately

$$\hat{\lambda} - (1.96)\sqrt{\frac{1}{i(\hat{\lambda})}} . \tag{9}$$

and the upper 95% normal confidence bound is

$$\hat{\lambda} + (1.96)\sqrt{\frac{1}{i(\hat{\lambda})}} . \tag{10}$$

For confidence intervals of the estimated mean survival time it is taken the reciprocals of the confidence bounds for λ.

4) Numerical Results

This section shows the numerical results using the two different approaches mentioned above.

Actual Times — Data Set 1 : Using actual lifetimes of data set 1 gives :

$\hat{\lambda}$: 0.0141848

95% confidence interval : (0.007632, 0.020738)

This translates into :

expected lifetime : 70.4978

95% confidence interval : (48.2209, 131.0312).

Compound Times — Data Set 1 : Using the minimun start time (7.6543 minutes) of the ten replications in data set 1 gives :

$\hat{\lambda}$: 0.0141052

95% confidence interval : (0.007589, 0.020622)

which translates into

expected lifetme : 70.8957

95% confidence interval : (48.4931, 131.7707).

5) Cross Validation

As indicated above, the model fit is checked by cross validation. The data set 2 was reserved for this purpose. Computing the value for $\hat{\lambda}$ using the actual lifetime from the second data set gives :

$\hat{\lambda}$: 0.0161057

expected lifetime : 62.0900.

The values fall well inside the confidencd bounds.

Using the compound start time (7.6477 minutes) from the data set 2 gives :

$\hat{\lambda}$: 0.0157424

expedted lifetime : 63.5226

95% confidence interval : (44.1660, 113.0837).

Computer program to calculate these numerical results is shown in appendix A, and input data form using the actual lifetime of data set 1 is attached in appendix B.

3. Conclusion

The assertions at the outset on the applicability of an exponential model proved to be valid. Once the appropriate model, time-censored, was identified the calcultions and fit checks were relatively straight forward. The simulation data (original data) used for analysis were consistent and easily cleaned once placed in APL (A Programming Language) format. (Ref. 5). This model is easily adaptable to other events recorded in the simulation history file.

The model developed provides excellent characterization of Blue helicopter attrition. Both versions of the exponential model provide us sufficient confidence in the results for the purpose of aggregated combat simulation. To expand the applicability of this model the next analysis should similarly examine varied scenario for the Blue helicopter system. The value of this model could play heavily in theater plans for employment of this system and the repair requirements of the system in combat.

References

1. Joint Analysis Directorate Orgainzation of the Joint Chiefs of Staff, "Catalog of Wargaming and Military Simulation Models", Washington D. C. 1986.
2. Sinha, S. K. "Reliability and Life Testing", John Wiley & Sons, New York, 1986.
3. Ross, Shelding M. "Introduction to Probability Models", Academic Press Inc., Orlando, Florida, 1985.
4. Jacob, P. "OA-4301 Class Note", U. S. Naval Postgraduate School, Monterey California 93943, March 1988.
5. IBM Corporation, "System-APL Language (Six Editon)", P. O. Box 50020, Programming Publishing, San Jose, California 95150, November 1983.

```
C                                   Appendix A
C
C    This program computes the M.L.E. of mean survival time and
C    its 95% confidence interval for a corresponding data.
C    (Input data formation is attached on appendix B)
C
      INTEGER     N
      PARAMETER  (N=80)
      INTEGER     NF, REP(N)
      REAL        CHECK, ST(N), FT(N), T(N), INF, LHAT, MST, LB, UB, MUB
     1            MLB, TT

      OPEN (31,FILE='ZST1.DAT',STATUS='OLD')
      OPEN (41,FILE='ZST1.OUT',STATUS='NEW')

      DATA        CHECK  /25.0/
      DATA        NF, TT /0, 0.0/

      DO 20 I = 1, N
         READ(FMT=10,UNIT=31) REP(I),ST(I),FT(I)
10       FORMAT(3X,I2,4X,F6.4,3X,F7.4)
         T(I) = FT(I) - ST(I)
         TT = TT + T(I)
         IF(FT(I) .LT. CHECK) NF = NF + 1
20    CONTINUE

      LHAT = NF / TT
      INF = NF / (LHAT**2)
      LB = LHAT - 1.96 * SQRT(1/INF)
      UB = LHAT + 1.96 * SQRT(1/INF)
      MST = 1 / LHAT
      MLB = 1 / UB
      MUB = 1 / LB

      WRITE(*,12) LHAT, MST, MLB, MUB
      WRITE(41,12) LHAT, MST, MLB, MUB
12    FORMAT(/3X,'***<< OUTPUT FOR MODEL DEVELOPMENT DATA >>***',
     1//10X,'EXPONENTIAL RATE    ===> ',F11.7,/10X,'MEAN SURVIVAL ',
     1'TIME ===> ',F8.4,//10X,'CONFIDENCE INTERVAL(95%)',
     1/15X,'LOWER BOUND ===> ',F8.4,//15X,'UPPER BOUND ===> ',F8.4)
      WRITE(6,'(/)')
      STOP
      END
```

```
   ***<< OUTPUT FOR MODEL DEVELOPMENT DATA >>***

          EXPONENTIAL RATE    ===>   0.0141848
          MEAN SURVIVAL TIME ===>  70.4978

          CONFIDENCE INTERVAL(95%)
               LOWER BOUND ===>  48.2209

               UPPER BOUND ===> 131.0312
```

Appendix B

Rep. No.	Start Time	Finish Time	Helo No.
21	7.7168	25.0000	101
21	7.7168	25.0000	102
21	7.7168	25.0000	103
21	7.7168	25.0000	105
21	7.7168	25.0000	106
21	7.7168	25.0000	107
21	7.7168	22.6826	108
21	7.7168	25.0000	109
22	7.7195	25.0000	101
22	7.7195	25.0000	102
22	7.7195	8.4517	103
22	7.7195	25.0000	105
22	7.7195	25.0000	106
22	7.7195	25.0000	107
22	7.7195	25.0000	108
22	7.7195	25.0000	109
23	7.7434	25.0000	101
23	7.7434	25.0000	102
23	7.7434	25.0000	103
23	7.7434	25.0000	105
23	7.7434	20.1824	106
23	7.7434	19.4529	107
23	7.7434	24.2051	108
23	7.7434	25.0000	109
24	7.6963	25.0000	101
24	7.6963	25.0000	102
24	7.6963	25.0000	103
24	7.6963	25.0000	105
24	7.6963	24.5557	106
24	7.6963	25.0000	107
24	7.6963	23.7117	108
24	7.6963	8.4536	109
25	7.6543	13.0762	101
25	7.6543	25.0000	102
25	7.6543	25.0000	103
25	7.6543	25.0000	105
25	7.6543	25.0000	106
25	7.6543	25.0000	107
25	7.6543	21.8748	108
25	7.6543	8.2974	109
26	7.7178	25.0000	101
26	7.7178	25.0000	102
26	7.7178	25.0000	103
26	7.7178	25.0000	105
26	7.7178	25.0000	106
26	7.7178	25.0000	107
26	7.7178	25.0000	108
26	7.7178	25.0000	109
27	7.7424	25.0000	101
27	7.7424	25.0000	102
27	7.7424	25.0000	103
27	7.7424	25.0000	105
27	7.7424	25.0000	106
27	7.7424	25.0000	107
27	7.7424	25.0000	108
27	7.7424	25.0000	109
28	7.6709	13.0339	101
28	7.6709	25.0000	102
28	7.6709	25.0000	103
28	7.6709	25.0000	105
28	7.6709	25.0000	106
28	7.6709	19.1589	107
28	7.6709	23.0969	108
28	7.6709	25.0000	109
29	8.0227	25.0000	101
29	8.0227	25.0000	102
29	8.0227	25.0000	103
29	8.0227	25.0000	105
29	8.0227	25.0000	106
29	8.0227	25.0000	107
29	8.0227	25.0000	108
29	8.0227	25.0000	109
30	7.7542	24.4514	101
30	7.7542	25.0000	102
30	7.7542	25.0000	103
30	7.7542	25.0000	105
30	7.7542	19.9436	106
30	7.7542	21.0972	107
30	7.7542	22.7410	108
30	7.7542	25.0000	109