Searching for an optimal design of Paper Helicopter

Smriti Singh May 3, 2021



### **Contents**

1.	Problem Description	3
2.	Introduction to Design of Experiment.	4
3.	Methodology	5
4.	Experiment	7
5.	Results	.12
6.	Validation	13
7.	Recommendation	14
A.	Appendix	
	A1. R code and output	15
	A2. Helicopter photo	45

### **Problem Description**

The goal of this project is to use Plackett-Burman Design and Hamada-Wu method to maximize the flight time of a paper helicopter. We implement design of experiment concepts and strategies to explore and develop a relationship between the independent and response variables and find the levels of variables of factors that maximize the response variable i.e., flight time.

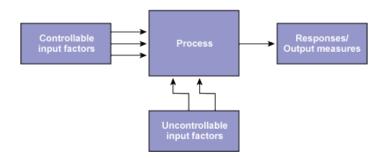
We begin the experiment by using the pattern provided by Prof. Rui Tuo to build our helicopter. It is an easy pattern to adjust and replicate. We are also provided the factors and levels that are identified to affect the flight time of our helicopter.

In order to determine the vital few factors that affect the flight the most, we perform a screening experiment using the Plackett- Burman design and stepwise regression under Hamada- Wu strategy to eliminate the least important factors and reach the factor effects to test the helicopter for maximum flight time. In order to validate our conclusion, we perform a final validation experiment with our chosen factors and their optimal levels to ascertain our experiment successful.

All data processing and analysis has been carried out with the help of Microsoft Excel and Rstudio.

#### **Introduction to Design of Experiments**

Design of experiments (DoE) is the planning, conducting, analyzing, and interpreting a test to determine the significant impact between factors affecting a process or response. It is a way to evaluate the cause and effects relationship, in simpler words. Response is the measurable outcomes or assessable characteristics of the experiment. The purpose of an experiment could be to maximize, minimize or target a response. Factors are the inputs to a process which can either be controllable or uncontrollable. Controllable factors are the ones that can be adjusted during the course of the experiment. Uncontrollable factors are the parameters that cannot be changed. They are the cause for variability or noise in an experiment. There are multiple factors that affect the experiment, only a few of them are vital.



The effect of uncontrollable factors can be minimized by utilizing the key concepts of DoE such as *randomization*, *replication* and *blocking*. Randomization ensures elimination of unwanted bias and distribute the variation throughout the experiment. Replication is introduced to obtain a reliability and precision in the experiment. Higher the number of observations, lower the variance. Blocking is performed to ensure homogeneity in the experimental units for increased efficiency.

Regression analysis is a statistical method carried out to examine the influence of *factors* or *predictors* or *independent* variables on *response* or *dependent* variable. It tests the *null hypothesis* that there is no relation between the predictors and response variables. This analysis forms a model with the predictors and response variable and the null hypothesis tests the goodness of fit of the model to the given data. It helps us identify the significance of the factors involved. This model can imply *correlation* between variables but may not necessarily imply *causation*.

Multiple factors can be manipulated simultaneously to study the important interactions between the factors which might be overlooked when performing an experiment analyzing one factor at a time. *Full factorial designs* are performed to study all possible combinations of factor effects and *fractional factorial designs* for a portion of all possible combinations. In designing experiments, the foundation of designs is factorial designs. The factorial designs work efficiently in order to assess the effect of several factors on a response. A 2<sup>k</sup> design is a full factorial design with k factors, each coded at two levels-low (-) and high (+). These levels are chosen to indicate a range of possible values for each factor. The decision of choosing the design to conduct an experiment can be driven by various factors such as interpretability, resource limitations, logistical inconveniency and expense. There is "no free lunch" for run size economy. There is often a tradeoff involved while choosing a suitable and feasible design for the experiment.

### Methodology

#### General scheme:

It is proposed here the experiment be designed in such a way that large number of factor levels combinations can be studied in smallest number of experimental runs. One such design is the *Plackett-Burman design*, which is mainly used to identify few important main effects. A *main effect* is the effect an independent variable has on the dependent variable, ignoring the effects of other independent variables. When the effect of one independent variable on the dependent variable depends on the level of other independent variable(s), it is called an *interaction effect*.

There are three fundamental guiding principles for variable selection in an experiment:

- i. *Effect hierarchy*:
  - Lower order effects are more important than higher order effects
  - Effects of same order are equally important
- ii. *Effect heredity*:
  - In order for an interaction to be significant, at least one of its parent factors should be significant
- iii. *Effect sparsity*:
  - The number of relatively important factors in an experiment is small

#### Aliasing of effects

As mentioned earlier, a fractional factorial design does not consider all possible combinations of factor levels. *Aliasing* or *confounding* occurs in such designs. When the estimate of an effect includes the influence of one or more effects, the effects are said to be *aliased*. As a result, these aliased effects cannot be distinguished from each other for estimation and are usually a linear combination of effects.

#### Design resolution

Design resolution indicates the degree of aliasing in a fractional factorial design. Resolution III designs are when no main effects are aliased with other main effects, but they are aliased with 2- factor interactions. Resolution IV design are where no main effects are aliased with other main effects or two factor interactions but are aliased with 3- factor interactions. Resolution IV designs are main effects *clear* design.

#### Choice of Design

Orthogonality is a property of the design that ensures independent estimations of parameters. This is essential to design of experiments as it makes the estimates of significant effects estimable. An optimal design reduces the experimentation cost, allows estimation of parameters with fewer runs, accommodates multiple factor levels and is practically feasible to conduct.

#### Plackett- Burman design

PB designs are non-regular screening design plans to identify the few significant factors from a list of many potential factors in an experiment when the complete knowledge of the system is unavailable. PB designs are usually Resolution III designs and hence, the main effects are heavily confounded with 2 factor interactions, PB designs assume the 2 factor interactions to negligible and are therefore used to study main effects. For a 12 run PB design with 11 factors, for each factor, its main effect is confounded with 45 2- factor interactions making it difficult to interpret the significance of the interactions. This design is constructed to have orthogonality and are economical to construct with run size (multiple of 4). For a smaller number of observations, these designs are very efficient.

### Construction of design

For this experiment, we shall use an OA  $(12, 2^{11})$  design which can accommodate up to eleven factors. OA stands for orthogonal array. The design matrix for a 12- run PB design is as shown below:

Run	A	В	С	D	$\mathbf{E}$	F	G	Η	Ι	J	K
1	+	+	-	+	+	+	-	-	-	+	-
2	-	+	+	-	+	+	+	-	-	-	+
3	+	-	+	+	-	+	+	+	-	-	-
4	-	+	-	+	+	-	+	+	+	-	-
5	-	-	+	-	+	+	-	+	+	+	-
6	-	-	-	+	-	+	+	-	+	+	+
7	+	-	-	-	+	-	+	+	-	+	+
8	+	+	-	-	-	+	-	+	+	-	+
9	+	+	+	-	-	-	+	-	+	+	-
10	-	+	+	+	-	-	-	+	-	+	+
11	+	-	+	+	+	-	-	-	+	-	+
12	-	-	-	-	-	-	-	-	-	-	-

12 run Plackett Burman design matrix

This design matrix will be further customized to run our experiment in the next section.

### **Experiment**

### Design matrix for the experiment

In order to perform the experiment and identify the important factors, we will construct the design matrix considering the factors and levels provided in the problem statement which is shown below:

			Levels			
Factor	Symbol	Design symbol	Low (-) level	High (+) level		
Wing length	1	A	3 inches	4.5 inches		
Wing width	W	В	1.8 inches	2.4 inches		
Body length	L	С	3 inches	4.5 inches		
Body width	W	D	1.25 inches	2 inches		
Middle body length	d	Е	1 inch	1.5 inches		
Fold at tip	F	F	no	yes		

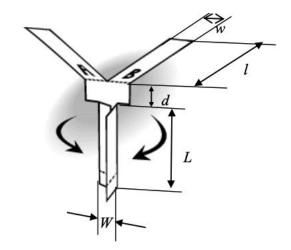
#### Protocol

In this experiment, we have to perform a 12-run experiment with six factors. Since the 12-run design allows for 11 factors to be considered, we shall randomly choose six columns from the matrix for our factors. This randomization was performed using R and columns 1, 3, 5, 8, 10, 11 were chosen to represent the six factors. For the experiment, the 12 runs were randomly carried out and the run order was again generated using R. The final design matrix for this experiment is as shown below:

Standard	Run	Factor									
order	order	A	В	С	D	Е	F				
1	5	+1	-1	+1	-1	+1	-1				
2	8	-1	+1	+1	-1	-1	+1				
3	3	+1	+1	-1	+1	-1	-1				
4	9	-1	-1	+1	+1	-1	-1				
5	7	-1	+1	+1	+1	+1	-1				
6	12	-1	-1	-1	-1	+1	+1				
7	11	+1	-1	+1	+1	+1	+1				
8	2	+1	-1	-1	+1	-1	+1				
9	10	+1	+1	-1	-1	+1	-1				
10	6	-1	+1	-1	+1	+1	+1				
11	4	+1	+1	+1	-1	-1	+1				
12	1	-1	-1	-1	-1	-1	-1				

The +1 represents the high level of the factors and -1 represents the low-level settings of the factors.

Based on the given factors, a template for the helicopter pattern was created and cut using the same quality of paper consistently.



For helicopter dropping, consideration about air circulation, dropping method and location and time recording was taken into account and an experimental protocol was set up to eliminate as many conditional factors as possible. The helicopter models for each run order were dropped from a height of 8 feet and its flight time was recorded on a stopwatch. If a helicopter model experienced any disturbance during the drop, that helicopter was redropped. This was done to ensure homogenous experimental environment.

#### Data Collection

Based on our protocol, 12 helicopter models were dropped from the same point, four times each (*for replication*), in the previously decided random order. The data collected from this experiment is shown below. The mean of the observations and standard deviation, variance and log of same is also shown below:

Std	Run		Factor Flight time						Mean	Std.	Var	ln s <sup>2</sup>			
order	order	Α	В	C	D	Е	F	1	2	3	4	Mean	dev	$s^2$ in s	
1	5	+1	-1	+1	-1	+1	-1	2.430	2.440	2.360	2.360	2.398	0.043	0.002	-6.270
2	8	-1	+1	+1	-1	-1	+1	2.010	1.980	1.960	1.950	1.975	0.026	0.001	-7.264
3	3	+1	+1	-1	+1	-1	-1	2.510	2.480	2.560	2.590	2.535	0.049	0.002	-6.018
4	9	-1	-1	+1	+1	-1	-1	2.130	1.830	2.060	2.010	2.008	0.128	0.016	-4.109
5	7	-1	+1	+1	+1	+1	-1	1.310	1.630	1.760	1.500	1.550	0.192	0.037	-3.300
6	12	-1	-1	-1	-1	+1	+1	2.080	2.160	2.080	2.180	2.125	0.053	0.003	-5.890
7	11	+1	-1	+1	+1	+1	+1	2.290	2.350	2.430	2.310	2.345	0.062	0.004	-5.564
8	2	+1	-1	-1	+1	-1	+1	2.110	1.890	2.210	1.680	1.973	0.236	0.056	-2.884
9	10	+1	+1	-1	-1	+1	-1	2.510	2.260	2.300	2.410	2.370	0.113	0.013	-4.364
10	6	-1	+1	-1	+1	+1	+1	2.080	1.960	2.160	2.060	2.065	0.082	0.007	-4.996
11	4	+1	+1	+1	-1	-1	+1	2.040	2.030	2.100	2.260	2.108	0.106	0.011	-4.484
12	1	-1	-1	-1	-1	-1	-1	1.810	1.900	2.090	1.910	1.928	0.117	0.014	-4.286

### Data Analysis

The primary goal here is to identify and quantify the influence of key factors for process improvement, which in turn will enable us to predict an optimal response under certain operating conditions.

### Effects Analysis

After the data points were collected, firstly, main effects significance was identified using *half normal plots*. These plots are graphical tools to help assess factors based on their estimated effects. Absolute values of estimated effects are ordered are plotted on the vertical axis and half normal distribution on the horizontal axis. All estimated effects can be obtained from the following form:

estimated effect = 
$$Y^-$$
 (+) -  $Y^-$  (-)

where Y (+) is the average of all responses for which the factor takes a '+' value and Y (-) is the average of responses for which the factor takes on a '-' value. These effects tend to follow a normal distribution. Normal probability plots are linear if normal. For factors that are unimportant, the absolute difference in '+' and '-' averages would be really small and therefore near linear while that of important factors would off the line. In a half normal plot, the effects on the extreme right are deemed statistically significant.

In this experiment, we plot the half normal for main effects only as interaction effects cannot be computed directly due to complex aliasing present. There are two types of half normal plots in this design:

- Location effect plot— to examine the relationship between the mean of the response and the factors
- Dispersion effect plot- to examine the relationship between the standard deviation of the replicate responses and the factors

The main effects (ordered absolute values) calculated for location and dispersion model are shown below:

Factor	Effect
A	0.381
C	0.137
В	0.063
D	0.037
Е	0.020
F	0.002

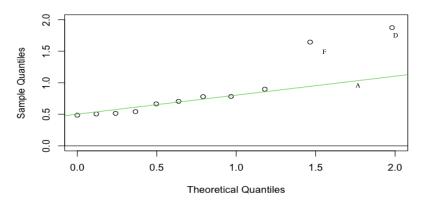
Location effect

Factor	Effect
D	0.898
F	0.505
С	0.376
В	0.188
Е	0.174
A	0.006

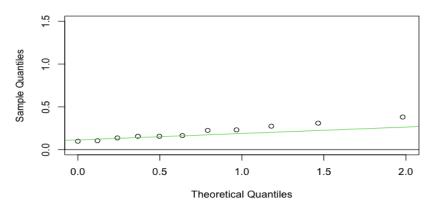
Dispersion effect

These main effects have been plotted on a half normal distribution plot and significant effects as per the location model can be identified.

#### **Half-Normal Dispersion Effect Plot**



#### **Half-Normal Location Effect Plot**



As per the location effect half normal plot, factor A (i.e., wing length) is the most significant factor that affects the response (i.e., flight time) in this experiment.

The PB design forces the factors to be partially aliased with interactions. These partial correlations can be taken advantage of to identify significant interactions. A method suggested by Hamada & Wu introduces an algorithm that identifies the significant factors while implementing the effect heredity and sparsity principle, without the requirement of *a priori* knowledge of the true model. It reduces the computation time and complexity while eliminating the need for searching all possible combinations of models.

This method called the *Hamada-Wu method*, uses the stepwise selection method and is performed in the following manner:

1. For each factor A, a model with A and all possible two-factor interactions (AB, AC, AD, etc.) involving A is considered. A stepwise regression is applied to select a model and record its significant effects. This process is repeated for all the factors in the experiment. Out of the 6 selected models, the best one is chosen.

- 2. The stepwise regression procedure is now performed to select significant effects from a model that consists of all the main effects and the two-factor interactions from the best model in the previous step.
- 3. Following the effect heredity principle, consider all the effects identified in the previous step, as well as the two-factor interactions that have at least one parent factor identified as significant in the previous step. The stepwise regression procedure is used again to determine the significant effects.
- 4. Using the model obtained in the previous step, an iteration between Steps 2 and 3 are performed until there are no further changes in effects identified as significant.

On implementing this algorithm on our response-factor model in R (code attached in Appendix), we can gather that the factor A (wing length) is the most significant factor in this paper helicopter experiment.

#### **Results**

In this experiment, our goal was to identify the significant factors that affect the response-flight time of the helicopter and to propose the optimal factor level settings for maximum flight time. Now that we concluded from the experiment that the factor A (wing length) has the most significance, all other factors and their levels can be chosen to meet our goal of maximum flight, depending on the economical and logistical constraints, if any.

The response, flight time, can be represented by the following regression equation:

$$y = \alpha + \beta_A x_A$$

where  $\alpha$  is the mean of responses (intercept),  $\beta$  is the effect of the individual factors and x represents the factor level settings.

For maximum value of response variable y, factor A should be set at high. The level setting of all other factors does not affect our helicopter model substantially and hence, can be set at high or low.

Calculating the value of y for factor A set at level +1 (high) results in flight time,

$$y = 2.115 + 0.19 = 2.305$$

Here, our model predicts the flight time of 2.305 seconds.

#### Validation

Now that we have the expected values of response variable, flight time for factor A level set at high and all other factors set low/ high, we can perform a confirmatory experiment to validate our results.

Two helicopters were manufactured for the factor levels not used in the model building process as shown below, dropped from the same drop point and flight time was recorded in similar experimental environment. The confirmatory experiment results are as follows:

Random		Factor					Flight time				Std.	Var	1-a a <sup>2</sup>	
order	Α	В	C	D	Е	F	1	2	3	4	n	dev	$s^2$	ln s <sup>2</sup>
1	. 1	+	+	. 1	+	+	2.41	2 31	2.29	2.26	2.31	0.06	0.00	-
1	+1	1	1	+1	1	1	2.41	2.31	2.29	2.20	8	5	4	5.467
2	. 1	1	1	1	1	1	2.47	266	2.46	2.54	2.53	0.09	0.00	-
2	+1	-1	-1	-1	-1	-1	2.47	2.66	2.46	2.54	3	2	8	4.769

This validation experiment gives a mean response of 2.533 secs with a standard deviation of 0.092 for factor A set high and all other factors set low and a mean response of 2.318 secs with a standard deviation of 0.065 for factor A set high and all other factors also set high.

Comparing these means with our predicted value of flight time (fit of experimental output of response from the model) obtained above, we can conclude that our analysis result is accurate and that irrespective of factors such as wing width, body length, body width, etc., the flight time is majorly affected by wing length. As we can see the percentage error between the maximum flight time obtained in validation experiment and the predicted value is 9.89%, which is relatively a small number.

#### Recommendation

Based on the experiment and analysis performed, it is recommended that the factor *wing length* be set at high level for maximum flight time. All other factors are not significantly important and therefore, can be set at low level in order to minimize the cost and improve the process efficiency in the construction of the paper helicopter.

Optimal factor level settings for maximum flight time:

Factor	Crymbal	Degian gymbol	Optimal Lo	evels
Factor	Symbol	Design symbol	Uncoded	Coded
Wing length	1	A	4.5 inches	High
Wing width	W	В	1.8 inches	Low
Body length	L	С	3 inches	Low
Body width	W	D	1.25 inches	Low
Middle body length	d	Е	1 inch	Low
Fold at tip	F	F	no	Low

This could be a starting point for further analyses to be considered to inspect other factors and de-aliased interactions that affect the response variable in the experiment.

#### **APPENDIX**

### *R- code and output:*

```
#randomizing the experiment
set.seed(23)
runseq <- 1:12
randomseq <- sample(runseq)</pre>
randomseq
set.seed(23)
factorcol <- c("1","2","3","4","5","6","7","8","9","10","11")
randomfactorcol <- sort(sample(factorcol,6))
randomfactorcol
Output:
>randomseq
[1] 12 8 3 11 1 10 5 2 4 9 7 6
> randomfactorcol
[1] "1" "10" "11" "3" "5" "8"
#halfnormal plot
library(readxl)
loc = read excel("locationeffect.xlsx")
attach(loc)
loceffect <- as.data.frame(loc)</pre>
str(loceffect)
gqnorm(loceffect$`Location effect`,xlim=c(0,2),ylim=c(0,1.5),main="Half-Normal Location Effect
Plot")
ggline(loceffect$`Location effect`, col=3)
abline(0,0)
disp = read excel("dispersioneffect.xlsx")
attach(disp)
dispeffect <- as.data.frame(disp)</pre>
str(dispeffect)
gqnorm(dispeffect$`Dispersion effect`,xlim=c(0,2),ylim=c(0,2),main="Half-Normal Dispersion Effect
Plot")
ggline(dispeffect$`Dispersion effect`, col=3)
abline(0,0)
Output: Plots shown in the main report.
#reading the data
library(readxl)
D = read_excel("regdata.xlsx")
attach(D)
#analyzing the data
names(D)
dim(D)
```

```
str(D)
data <- as.data.frame(D)
str(data)
plot(data)
cor(data)
#Hamada- Wu stepwise regression method
##Step 1
mA < -lm(Y1 \sim A + A:B + A:C + A:D + A:E + A:F, data = data)
kA <- ols_step_both_p(mA, pent = 0.1, prem = 0.3, details =TRUE)
kA
mB < -lm(Y1 \sim B + B:A + B:C + B:D + B:E + B:F, data = data)
kB <- ols_step_both_p(mB, pent = 0.1, prem = 0.3, details =TRUE)
mC < -lm(Y1 \sim C + C:A + C:B + C:D + C:E + C:F, data = data)
kC <- ols_step_both_p(mC, pent = 0.1, prem = 0.3, details =TRUE)
mD < -lm(Y1 \sim D + D:A + D:B + D:C + D:E + D:F, data = data)
kD < -ols step both p(mD, pent = 0.1, prem = 0.3, details = TRUE)
kD
mE < -lm(Y1 \sim E + E:A + E:B + E:C + E:D + E:F, data = data)
kE <- ols_step_both_p(mE, pent = 0.1, prem = 0.3, details =TRUE)
mF < -lm(Y1 \sim F + F:A + F:B + F:C + F:D + F:E, data = data)
kF < - ols step both p(mF, pent = 0.1, prem = 0.3, details = TRUE)
kF
##Step 2
m2 < -lm(Y1 \sim A + B + C + C:B + D + E + F:A + F:D, data = data)
k2 < - ols step both p(m2, pent = 0.1, prem = 0.3, details =TRUE)
k2
##Step 3
m4 < -lm(Y1 \sim A + A:B + A:C + A:D + A:E + A:F, data = data)
k4 <- ols_step_both_p(m4, pent = 0.1, prem = 0.3, details =TRUE)
k4
Output:
Stepwise Selection Method
Candidate Terms:
1. A
2. A:B
3. A:C
```

4. A:D

5. A:E

6. A:F

We are selecting variables based on p value...

Stepwise Selection: Step 1

+A

### Model Summary

R	0.755	RMSE	0.181
R-Squared	0.570	Coef. Var	8.494
Adj. R-Square	d 0.527	MSE	0.033
Pred R-Square	d 0.381	MAE	0.127

RMSE: Root Mean Square Error

MSE: Mean Square Error MAE: Mean Absolute Error

#### ANOVA

	m of uares	DF	Mean Square	F	Sig.
Regression Residual Total	0.435 0.328 0.763	1 10 11	0.435 0.033	13.267	0.0045

### Parameter Estimates

model	Beta	Std. Error	Std. Be	eta t	Sig	lower	upper
(Intercept) A 0		0.267 0.070				0.586 0.099	

-----

No more variables to be added/removed.

Final Model Output

-----

Model	Summary

R	0.755	RMSE	0.181

R-Squared	0.570	Coef. Var	8.494
Adj. R-Squared	0.527	MSE	0.033
Pred R-Squared	0.381	MAE	0.127

RMSE: Root Mean Square Error

MSE: Mean Square Error MAE: Mean Absolute Error

#### **ANOVA**

Sum of Squares DF Mean Square F Sig.

Regression 0.435 1 0.435 13.267 0.0045 Residual 0.328 10 0.033 Total 0.763 11

Parameter Estimates

model	Beta	Std. Error	Std. Beta	t	Sig lo	ower u	pper
(Intercept)	1.180	0.267		4.427	0.001	0.586	 1.774
A	0.254	0.070	0.755	3.642	0.005	0.099	0.409

> kA

#### Stepwise Selection Summary

Added/ Adj.

Step Variable Removed R-Square R-Square C(p) AIC RMSE

1 A addition 0.570 0.527 -1.7080 -3.1432 0.1811

### Stepwise Selection Method

-----

#### Candidate Terms:

- 1. B
- 2. B:A
- 3. B:C
- 4. B:D
- 5. B:E
- 6. B:F

We are selecting variables based on p value...

Stepwise Selection: Step 1

+ B:A

### Model Summary

R	0.559	RMSE	0.229
R-Squared	0.313	Coef. Var	10.741
Adj. R-Squared	0.24	4 MSE	0.052
Pred R-Squared	0.06	MAE	0.154

RMSE: Root Mean Square Error

MSE: Mean Square Error MAE: Mean Absolute Error

#### **ANOVA**

	ım of			·	
Sq	uares	DF 	Mean Square	F	Sig.
Regression	0.239		0.239	4.55	0.0587
Residual	0.524	10	0.052		
Total	0.763	11			

Parameter Estimates

model	Beta	Std. Error	Std. Be	ta t	Sig	lower	upper
` '		0.275 0.034					

Stepwise Selection: Step 2

+ B

### Model Summary

R	0.784	<b>RMSE</b>	0.181	
R-Squared	0.614	Coef.	Var 8.4	480
Adj. R-Squared	0.5	29 MS	E 0	.033
Pred R-Squared	0.2	84 MA	AE (	0.127
-				

RMSE: Root Mean Square Error

MSE: Mean Square Error MAE: Mean Absolute Error

A	N	$\cap$	V	Α

~ .	um of uares	DF	Me	ean Square	F	Sig.
Regression Residual Total	0.469 0.294 0.763	9 11	2	0.234 0.033	7.17	0.0137

Parameter Estimates

model	Beta	Std. Error	Std. Be	 ta t	Sig	lower	upper
	).565	0.369 0.213 0.033	-0.673	-2.653	0.026	-1.047	-0.083

### Model Summary

R	0.784 R	MSE	0.181
R-Squared	0.614	Coef. Var	8.480
Adj. R-Squared	0.529	MSE	0.033
Pred R-Squared	0.284	MAE	0.127

·

RMSE: Root Mean Square Error

MSE: Mean Square Error MAE: Mean Absolute Error

### ANOVA

	um of uares	DF	Me	an Square	F	Sig.
Regression Residual Total	0.469 0.294 0.763	9 11	2	0.234 0.033	7.17	0.0137

### Parameter Estimates

m	odel	Beta	Std. Error	Std. Beta	t	Sig	lower	upper

20

(Intercept)	2.354	0.369		6.377	0.000	1.519	3.189
В	-0.565	0.213	-0.673	-2.653	0.026	-1.047	-0.083
B:A	0.123	0.033	0.948	3.738	0.005	0.048	0.197

Stepwise Selection: Step 3

x B:A

### Model Summary

R 0.126 RMSE 0.274
R-Squared 0.016 Coef. Var 12.853
Adj. R-Squared -0.083 MSE 0.075
Pred R-Squared -0.417 MAE 0.190

RMSE: Root Mean Square Error

MSE: Mean Square Error MAE: Mean Absolute Error

#### **ANOVA**

	um of uares	DF	Mean Square	F	Sig.
Regression Residual Total	0.012 0.751 0.763	10 11		0.16	0.6974

#### Parameter Estimates

model	Beta	Std. Error	Std. Be	ta t	Sig	lower	upper
(Intercept) B -0		0.559 0.264					

No more variables to be added/removed.

Final Model Output

-----

R	0.126 RN	<b>ISE</b>	0.274
R-Squared	0.016	Coef. Var	12.853
Adj. R-Squared	-0.083	MSE	0.075
Pred R-Squared	-0.417	MAE	0.190

-----

RMSE: Root Mean Square Error

MSE: Mean Square Error MAE: Mean Absolute Error

### ANOVA

	um of uares	DF N	Mean Square	F	Sig.
Regression Residual Total	0.012 0.751 0.763	1 10 11	0.012 0.075	0.16	0.6974

\_\_\_\_\_

#### Parameter Estimates

model	Beta	Std. Error	Std. Be	eta t	Sig	lower	upper
(Intercept) B -0		0.559 0.264					

> kB

### **Stepwise Selection Summary**

		Added/	Ad	j.				
Step	Varial	ole Remov	ed R-So	quare R	L-Square	C(p)	AIC	<b>RMSE</b>
1	B:A	addition	0.313	0.244	3.5510	2.4900	0.2290	
2	В	addition	0.614	0.529	0.4810	-2.4449	0.1808	
3	B:A	removal	0.016	-0.083	8.5410	6.7997	0.2740	)

Stepwise Selection Method

-----

Candidate Terms:

1. C

2. C:A

3. C:B

4. C:D

5. C:E

6. C:F

We are selecting variables based on p value...

Stepwise Selection: Step 1

+ C:B

### Model Summary

R	0.362 R	MSE	0.257
R-Squared	0.131	Coef. Var	12.076
Adj. R-Squared	0.044	MSE	0.066
Pred R-Squared	-0.304	MAE	0.198

RMSE: Root Mean Square Error

MSE: Mean Square Error MAE: Mean Absolute Error

### **ANOVA**

	um of uares	DF	Mean Square	F	Sig.
Regression Residual Total	0.100 0.663 0.763	10 11		1.511	0.2471

### Parameter Estimates

model	Beta	Std. Error	Std. Bet	a t	Sig	lower	upper
(Intercept) C:B		0.309 0.038					

Stepwise Selection: Step 2

+ C:A

Model Summary


R	0.726 RN	<b>MSE</b>	0.200
R-Squared	0.527	Coef. Var	9.388
Adj. R-Squared	0.422	MSE	0.040
Pred R-Squared	0.198	MAE	0.136

RMSE: Root Mean Square Error

MSE: Mean Square Error MAE: Mean Absolute Error

### ANOVA

	um of uares	DF	Mean	Square	F	Sig.
Regression Residual Total	0.402 0.361 0.763	9 11	2	0.201 0.040	5.023	0.0343

### Parameter Estimates

model	Beta	Std. Error	Std. Bet	a t	Sig	lower	upper
C:B	-0.095	0.259 0.035 0.017	-0.738	-2.765	0.022	-0.174	-0.017

### Model Summary

0.726 H	RMSE	0.200
0.527	Coef. Var	9.388
0.422	2 MSE	0.040
0.19	8 MAE	0.136
	0.527 1 0.422	0.527 Coef. Var 0.422 MSE

RMSE: Root Mean Square Error

MSE: Mean Square Error

### MAE: Mean Absolute Error

### ANOVA

~ .	um of uares	DF	Mean Square	F	Sig.
Regression Residual Total	0.402 0.361 0.763	9 11	0.201 0.040	5.023	0.0343

### Parameter Estimates

model	Beta	Std. Error	Std. Bet	a t	Sig	lower	upper
C:B	-0.095	0.259 0.035 0.017	-0.738	-2.765	0.022	-0.174	-0.017

Stepwise Selection: Step 3

+ C

### Model Summary

0.823	RN	<b>ASE</b>	0.175	
0.6	77	Coef. V	7ar 8.2	228
1 0	.556	MSE	0.	.031
1 0	.255	MAE	E C	.122
	0.6 I 0	0.677 1 0.556	0.677 Coef. V 0.556 MSE	0.677 Coef. Var 8.2 1 0.556 MSE 0.

RMSE: Root Mean Square Error MSE: Mean Square Error MAE: Mean Absolute Error

### ANOVA

~ .	um of uares	DF	Mean Square	F	Sig.
Regression				5.598	0.0230
Residual	0.246	8	0.031		
Total	0.763	11			

### Parameter Estimates

model	Beta	Std. Error	Std. Bet	 ta t 	Sig	lower	upper
C C:B	2.474 -0.268 -0.033 0.065	0.258 0.139 0.045 0.018	-0.798 -0.251	-1.928 -0.731		-0.589 -0.135	0.053 0.070

### Model Summary

R	0.823 RN	MSE	0.175
R-Squared	0.677	Coef. Var	8.228
Adj. R-Squared	0.556	MSE	0.031
Pred R-Squared	0.255	MAE	0.122

RMSE: Root Mean Square Error

MSE: Mean Square Error MAE: Mean Absolute Error

#### ANOVA

~ .	ım of uares	DF	Mean Square	F	Sig.
Regression Residual Total	0.517 0.246 0.763	8 11	0.172 0.031	5.598	0.0230

### Parameter Estimates

model	Beta	Std. Error	Std. Bet	a t	Sig	lower	upper
C C:B	-0.268	0.045	-0.798	-1.928 -0.731	0.090 0.486	-0.589 -0.135	0.053 0.070

\_\_\_\_\_\_

Stepwise Selection: Step 4

x C:A

### Model Summary

R	0.364 RN	364 RMSE	
R-Squared	0.133	Coef. Var	12.719
Adj. R-Squared	-0.060	MSE	0.074
Pred R-Squared	-0.553	MAE	0.198

RMSE: Root Mean Square Error

MSE: Mean Square Error MAE: Mean Absolute Error

### ANOVA

~	um of uares	DF	Mean Square	F	Sig.
Regression Residual Total	0.101 0.662 0.763	 9 11	2 0.051 0.074	0.688	0.5271

#### Parameter Estimates

model	Beta	Std. Error	Std. Bet	 ta t	Sig	lower	upper
C	0.021	0.399 0.177 0.068	0.063	0.120	0.907	-0.380	0.422

No more variables to be added/removed.

### Final Model Output

-----

### Model Summary

R	0.364 RN	MSE	0.271
R-Squared	0.133	Coef. Var	12.719
Adj. R-Squared	-0.060	MSE	0.074
Pred R-Squared	-0.553	MAE	0.198

RMSE: Root Mean Square Error

MSE: Mean Square Error MAE: Mean Absolute Error

### ANOVA

	um of uares	DF	Me	an Square	F	Sig.
Regression Residual Total	0.101 0.662 0.763	9 11	2	0.051 0.074	0.688	0.5271

\_\_\_\_\_

### **Parameter Estimates**

model	Beta	Std. Error	Std. Bet	a t	Sig	lower	upper
C	0.021	0.399 0.177 0.068	0.063	0.120	0.907	-0.380	0.422

-----

> kC

## Stepwise Selection Summary

		Added/	Ad	j.				
Step	Varial	ble Remove	ed R-Sq	juare R	-Square	C(p)	AIC	RMSE
1	C:B	addition	0.131	0.044	5.7340	5.3016	0.2575	
2	C:A	addition	0.527	0.422	1.4700	-0.0058	0.2002	
3	C	addition	0.677	0.556	1.1010	-2.5841	0.1754	
4	C:A	removal	0.133	-0.060	7.7120	7.2825	0.2712	2

\_\_\_\_\_\_

### Stepwise Selection Method

-----

### Candidate Terms:

- 1. D
- 2. D:A
- 3. D:B
- 4. D:C
- 5. D:E
- 6. D:F

We are selecting variables based on p value...

### Stepwise Selection: Step 1

#### + D:A

### Model Summary

0.464 RN	<b>MSE</b>	0.245
0.215	Coef. Var	11.478
0.137	MSE	0.060
-0.017	MAE	0.160
	0.215 0.137	0.215 Coef. Var 0.137 MSE

RMSE: Root Mean Square Error

MSE: Mean Square Error MAE: Mean Absolute Error

### ANOVA

	um of uares	DF	Mean Square	F	Sig.
Regression Residual Total	0.164 0.599 0.763	10 10 11		2.741	0.1288

#### Parameter Estimates

model	Beta	Std. Error	Std. Be	 ta t	Sig	lower	upper
(Intercept) D:A		0.239 0.038			0.000 0.129		

Stepwise Selection: Step 2

+D

### Model Summary

R	0.783 R	MSE	0.181
R-Squared	0.613	Coef. Var	8.495
Adj. R-Squared	0.527	MSE	0.033
Pred R-Squared	0.273	MAE	0.133

RMSE: Root Mean Square Error

MSE: Mean Square Error MAE: Mean Absolute Error

### **ANOVA**

	um of uares	DF	Mea	n Square	F	Sig.
Regression Residual Total	0.468 0.295 0.763	9 11	2	0.234 0.033	7.131	0.0139

### Parameter Estimates

model	Beta	Std. Error	Std. Bet	 ta t	Sig	lower	upper
	).638	0.233 0.210 0.042	-0.949	-3.043	0.014	-1.113	-0.164

### Model Summary

R	0.783 R	MSE	0.181
R-Squared	0.613	Coef. Var	8.495
Adj. R-Squared	0.527	MSE	0.033
Pred R-Squared	0.273	MAE	0.133

RMSE: Root Mean Square Error

MSE: Mean Square Error MAE: Mean Absolute Error

### ANOVA

~	um of Juares	DF	Mean Square	F	Sig.
Regression Residual Total	0.468 0.295 0.763	9 11	0.234 0.033	7.131	0.0139

### Parameter Estimates

model	Beta	Std. Error	Std. Bet	 ta t	Sig	lower	upper
	).638	0.233 0.210 0.042	-0.949	-3.043	0.014	-1.113	-0.164

Stepwise Selection: Step 3

x D:A

### Model Summary

R	0.073 RN	<b>ISE</b>	0.276
R-Squared	0.005	Coef. Var	12.922
Adj. R-Squared	-0.094	MSE	0.076
Pred R-Squared	-0.432	MAE	0.198

RMSE: Root Mean Square Error

MSE: Mean Square Error MAE: Mean Absolute Error

### **ANOVA**

	um of uares	DF	Mean Square	F	Sig.
Regression Residual Total	0.004 0.759 0.763	10 11		0.053	0.8223

### Parameter Estimates

model	Beta	Std. Error	Std. Be	ta t	Sig	lower	upper
(Intercept) D -0		0.354 0.212					

No more variables to be added/removed.

### Final Model Output

-----

### Model Summary

R	0.073 RN	<b>MSE</b>	0.276
R-Squared	0.005	Coef. Var	12.922
Adj. R-Squared	-0.094	MSE	0.076
Pred R-Squared	-0.432	MAE	0.198

RMSE: Root Mean Square Error

MSE: Mean Square Error MAE: Mean Absolute Error

### ANOVA

	um of quares	DF	Mean Square	F	Sig.
Regression Residual Total	0.004 0.759 0.763	10 11		0.053	0.8223

### Parameter Estimates

model	Beta	Std. Error	Std. Be	ta t	Sig	lower	upper
(Intercept) D -(		0.354 0.212					
> kD							

### **Stepwise Selection Summary**

Step		Added/ ole Remove	-	,	-Square	C(p)	AIC	RMSE
1 2 3	D:A D D:A	aaamon	0.613	0.527	0.1010	-2.4050	0.1811	

### Stepwise Selection Method

-----

### Candidate Terms:

- 1. E
- 2. E:A
- 3. E:B
- 4. E:C
- 5. E:D
- 6. E:F

We are selecting variables based on p value...

Stepwise Selection: Step 1

+ E:A

### Model Summary

R	0.578	<b>RMSE</b>	0.225	
R-Squared	$0.33^{2}$	4 Coef.	. Var 10.575	
Adj. R-Squared	0.2	.67 MS	0.051	
Pred R-Squared	0.2	200 MA	AE 0.108	3

RMSE: Root Mean Square Error

MSE: Mean Square Error MAE: Mean Absolute Error

#### ANOVA

	um of uares	DF	Mean Square	F	Sig.
Regression Residual Total	0.255 0.508 0.763	1( 11		5.01	0.0491

### Parameter Estimates

model	Beta	Std. Error	Std. Be	ta t	Sig	lower	upper
` '		0.237 0.049			0.000 0.049		

Stepwise Selection: Step 2

+E

### Model Summary

				-
R	0.771	<b>RMSE</b>	0.185	
R-Squared	0.595	5 Coef	. Var 8.69	€2
Adj. R-Squared	0.5	05 MS	SE 0.0	134
Pred R-Squared	0.2	257 MA	AE 0.	116

RMSE: Root Mean Square Error

MSE: Mean Square Error MAE: Mean Absolute Error

#### **ANOVA**

~ .	um of uares	DF	Mea	n Square	F	Sig.
Regression Residual Total	0.454 0.309 0.763	9 11	2	0.227 0.034	6.608	0.0171

### Parameter Estimates

model	Beta	Std. Error	Std. Be	ta t	Sig	lower	upper
	0.722	0.273 0.300 0.056	-0.716	-2.409	0.039	-1.400	-0.044

\_\_\_\_\_

### Model Summary

R	0.771 RN	MSE	0.185
R-Squared	0.595	Coef. Var	8.692
Adj. R-Squared	0.505	MSE	0.034
Pred R-Squared	0.257	MAE	0.116

RMSE: Root Mean Square Error

MSE: Mean Square Error MAE: Mean Absolute Error

#### **ANOVA**

~ -	um of uares	DF	Mean Square	F	Sig.
Regression Residual Total	0.454 0.309 0.763	9 11	2 0.227 0.034	6.608	0.0171

Parameter Estimates

model	Beta	Std. Error	Std. Be	 ta t	Sig	lower	upper
	).722	0.273 0.300 0.056	-0.716	-2.409	0.039	-1.400	-0.044

\_\_\_\_\_

Stepwise Selection: Step 3

x E:A

### Model Summary

R	0.040 RM	MSE	0.276
R-Squared	0.002	Coef. Var	12.946
Adj. R-Squared	-0.098	MSE	0.076
Pred R-Squared	-0.438	MAE	0.193

RMSE: Root Mean Square Error

MSE: Mean Square Error MAE: Mean Absolute Error

#### ANOVA

~ .	um of uares	DF	Mean Square	F	Sig.
Regression Residual Total	0.001 0.762 0.763	10 11		0.016	0.9026

### Parameter Estimates

model	Beta	Std. Error	Std. B	eta t	Sig	lower	upper
(Intercept) E 0		0.406 0.319					

No more variables to be added/removed.

### Final Model Output

-----

### Model Summary

R	0.040 RN	<b>ISE</b>	0.276
R-Squared	0.002	Coef. Var	12.946
Adj. R-Squared	-0.098	MSE	0.076
Pred R-Squared	-0.438	MAE	0.193

RMSE: Root Mean Square Error

MSE: Mean Square Error MAE: Mean Absolute Error

### **ANOVA**

	um of uares	DF	Mean Square	F	Sig.
Regression Residual Total	0.001 0.762 0.763	10 11		0.016	0.9026

### Parameter Estimates

model	Beta	Std. Error	Std. B	eta t	Sig	lower	upper
(Intercept) E 0.		0.406 0.319				1.177 -0.670	

> kE

### **Stepwise Selection Summary**


		Added/	Ad	lj.				
Step	Varia	ble Remov	ed R-So	quare R	A-Square	C(p)	AIC	RMSE
1	E:A	addition	0.334	0.267	5.1020	2.1166	0.2255	j .
2	E	addition	0.595	0.505	1.9670	-1.8529	0.1853	
3	E:A	removal	0.002	-0.098	11.6350	0 6.971:	0.276	50

### Stepwise Selection Method

-----

### Candidate Terms:

- 1. F
- 2. F:A
- 3. F:B
- 4. F:C
- 5. F:D
- 6. F:E

We are selecting variables based on p value...

Stepwise Selection: Step 1

+ F:A

### Model Summary

R	0.084 RN	084 RMSE	
R-Squared	0.007	Coef. Var	12.910
Adj. R-Squared	-0.092	MSE	0.076
Pred R-Squared	-0.415	MAE	0.192

RMSE: Root Mean Square Error

MSE: Mean Square Error MAE: Mean Absolute Error

#### **ANOVA**

	um of uares	 DF	Mean	Square	F	Sig.
Regression			 1		0.071	0.7946
Residual	0.758	10		0.076		
Total	0.763	11				

### Parameter Estimates

model	Beta	Std. Error	Std. Be	ta t	Sig	lower	upper
(Intercept) F:A		0.079 0.021					

Stepwise Selection: Step 2

+ F:D

### Model Summary

R	0.538 RN	<b>MSE</b>	0.245
R-Squared	0.290	Coef. Var	11.509
Adj. R-Squared	0.132	MSE	0.060
Pred R-Squared	-0.318	MAE	0.168

RMSE: Root Mean Square Error

MSE: Mean Square Error MAE: Mean Absolute Error

#### ANOVA

	um of uares	DF	Mean	Square	F	Sig.
Regression	0.221		2	0.111	1.837	0.2143
Residual	0.542	9	(	0.060		
Total	0.763	11				

### Parameter Estimates

model	Beta	Std. Error	Std. Be	ta t	Sig	lower	upper
F:A	-0.119	0.071 0.063 0.144	-1.806	-1.897	0.090	-0.261	0.023

### Model Summary

R	0.538 R	MSE	0.245
R-Squared	0.290	Coef. Var	11.509
Adi. R-Squared	0.132	MSE	0.060

Pred R-Squared -0.318 MAE 0.168

-----

RMSE: Root Mean Square Error

MSE: Mean Square Error MAE: Mean Absolute Error

#### ANOVA

Sum of Squares DF Mean Square F Sig.

Regression 0.221 2 0.111 1.837 0.2143 Residual 0.542 9 0.060 Total 0.763 11

\_\_\_\_\_

#### Parameter Estimates

model	Beta	Std. Error	Std. Bet	ta t	Sig	lower	upper
	2.132 -0.119 0.272	0.071 0.063 0.144	-1.806	-1.897	0.090	-0.261	0.023

No more variables to be added/removed.

#### Final Model Output

-----

#### Model Summary

R 0.538 RMSE 0.245
R-Squared 0.290 Coef. Var 11.509
Adj. R-Squared 0.132 MSE 0.060
Pred R-Squared -0.318 MAE 0.168

RMSE: Root Mean Square Error

MSE: Mean Square Error MAE: Mean Absolute Error

#### ANOVA

 Sum of Squares
 DF Mean Square
 F Sig.

 Regression
 0.221
 2
 0.111
 1.837
 0.2143

Residual 0.542 9 0.060 Total 0.763 11

\_\_\_\_\_

### **Parameter Estimates**

model	Beta	Std. Error	Std. Be	ta t	Sig	lower	upper
F:A	-0.119	0.071 0.063 0.144	-1.806	-1.897	0.090	-0.261	0.023

> kF

### Stepwise Selection Summary

	A	Added/	Ad	j.				
Step	Variab	le Remove	ed R-So	quare R	-Square	C(p)	AIC	RMSE
1	F:A	addition	0.007	-0.092	1.2430	6.9049	0.2753	
2	F:D	addition	0.290	0.132	0.6100	4.8828	0.2454	

.....

> k2 <- ols\_step\_both\_p(m2, pent = 0.1, prem = 0.3, details =TRUE) Stepwise Selection Method

-----

### Candidate Terms:

- 1. A
- 2. B
- 3. C
- 4. D
- 5. E
- 6. B:C
- 7. A:F
- 8. D:F

We are selecting variables based on p value...

Stepwise Selection: Step 1

+A

### Model Summary

R	0.755 RM	MSE	0.181
R-Squared	0.570	Coef. Var	8.494
Adj. R-Squared	0.527	MSE	0.033
Pred R-Squared	1 0.381	MAE	0.127

RMSE: Root Mean Square Error

MSE: Mean Square Error MAE: Mean Absolute Error

### ANOVA

Sum of Squares DF Mean Square F Sig.

Regression 0.435 1 0.435 13.267 0.0045
Residual 0.328 10 0.033
Total 0.763 11

#### Parameter Estimates

model	Beta	Std. Error	Std. Bo	eta t	Sig	lower	upper
` '		0.267 0.070					

No more variables to be added/removed.

### Final Model Output

-----

### Model Summary

R	0.755 RN	MSE	0.181
R-Squared	0.570	Coef. Var	8.494
Adj. R-Squared	0.527	MSE	0.033
Pred R-Squared	0.381	MAE	0.127

RMSE: Root Mean Square Error

MSE: Mean Square Error MAE: Mean Absolute Error

#### ANOVA

	um of uares	DF	Mean Square	F	Sig.
Regression Residual Total	0.435 0.328 0.763	10 11		13.267	0.0045

\_\_\_\_\_

#### Parameter Estimates

model	Beta	Std. Error	Std. Bo	eta t	Sig	lower	upper
(Intercept) A 0		0.267 0.070					
< 1-2							

> k2

### Stepwise Selection Summary

Step	Varia	Added/ ble Remov		3	R-Square	C(p)	AIC	RMSE
1	A	addition	0.570	0.52	7 5.6130	-3.1432	0.1811	

> k4 <- ols\_step\_both\_p(m4, pent = 0.1, prem = 0.3, details =TRUE) Stepwise Selection Method

-----

#### Candidate Terms:

- 1. A
- 2. A:B
- 3. A:C
- 4. A:D
- 5. A:E
- 6. A:F

We are selecting variables based on p value...

Stepwise Selection: Step 1

+A

### Model Summary

R	0.755 R	MSE	0.181
R-Squared	0.570	Coef. Var	8.494
Adj. R-Squared	0.527	MSE	0.033
Pred R-Squared	0.381	MAE	0.127

RMSE: Root Mean Square Error

MSE: Mean Square Error MAE: Mean Absolute Error

### ANOVA

	um of uares	DF	Mean Square	F	Sig.
Regression	0.435		0.435	13.267	0.0045
Residual	0.328	10	0.033		
Total	0.763	11			

#### Parameter Estimates

model	Beta	Std. Error	Std. Be	eta t	Sig	lower	upper
(Intercept) A 0		0.267 0.070					

No more variables to be added/removed.

### Final Model Output

-----

# Model Summary

R	0.755	RMSE	0.181
R-Squared	0.570	Coef. Var	8.494
Adj. R-Squared	0.52	7 MSE	0.033
Pred R-Squared	0.38	MAE	0.127

RMSE: Root Mean Square Error

MSE: Mean Square Error MAE: Mean Absolute Error

٨	N.T		<b>T</b> 7	٨
Α	ΙN	U	V	А

	um of uares	DF	Mean Square	F	Sig.
Regression Residual Total	0.435 0.328 0.763	10 11		13.267	0.0045

### Parameter Estimates

model	Beta	Std. Error	Std. Bo	eta t	Sig	lower	upper
(Intercept) A 0		0.267 0.070					
> k4							

# Stepwise Selection Summary

Step	Varia	Added/ able Remov		J	R-Square	C(p)	AIC	RMSE
1	A	addition	0.570	0.527	7 -1.7080	-3.1432	0.1811	

