Design of Experiments – Golf Toy

ENGI 9516 – Similitude, Modelling, and Data Analysis

Golf Toy Team Project

Team B

- Greg Browne
- Keith Downey
- Carolyn Duncan
- Farzad Eskandari

April 3, 2008

Table of Contents

Table of Figures

Introduction

The DOE Golfer toy was invented by Dr. Leonard Lye to teach DOE (Design of Experiments) to students. The aim was to find a statistical model that predicted the operating conditions required to shoot a golf ball to within 5 cm of any target. The DOE Golfer toy is capable of changing the length of club, the angle of swing and the weight of the club. Assuming that the machine was to be used on the same ground with the same ball, it was possible to eliminate 2 factors – the type of ball and the type of ground.

A golf tournament was organized and the teams were given 2 weeks to practice and develop their model(s). The tournament consisted of two parts – first the teams were required to shoot the ball toward a line and maintain a tolerance of 2 cm. Secondly, 3 holes were determined and each team proceeded to shoot their ball toward the hole. For each group, the number of strokes to accomplish each hole was accumulated and, at the end, the team with the least number of strokes was the winner.

Methodology

Factors and Responses

In order to accurately shoot the ball the required distance using the golfer toy, a number of factors had to be considered. The four factors investigated were pull back angle of the shooting arm, length of the shooting arm, weights attached to the shooting arm, and direction that the golfer is facing. The angle, length of shooting arm and weight were identified for investigation because they all directly affect the force applied to, and resultant distance travelled by, the ball. Direction and location of the holes were considered because of the significant contours in the floor of the faculty lounge. These contours could

potentially affect both distance and deviation of the ball when struck. Factors and their chosen levels are listed in Table 1. Levels for all factors were chosen in an attempt to maximize the types of shots that the model could accurately predict. The angle and direction were input into the design as numeric factors, while both length of arm and number of weights were categorical factors.

Factor	Range
Angle (deg)	10-70
Length of Arm	1-3
Number of Weights	1-3
Direction (deg)	0-240

Table 1 - Ranges of Design Factors

Two response variables were determined to be important in the design of the putting model; most importantly was the distance the ball traveled. Deviation was also deemed to be important because of the aforementioned contours in the faculty lounge floor. When attempting to make a shot with these contours coming into play, it was equally important to not only predict how far it will deviate from its intended path, but also accurately predict the deviation (both measured in cm).

RSM

To develop a predictive model for the golfer toy, a central composite face-centred design was used. There are a number of reasons why this type of model was chosen over others. It was chosen over other types of CCD (central composite design) models because it did not require odd points that are not possible without significantly altering the golfer toy. Although this design is not rotatable, it is the only CCD that is viable for this experiment. Additionally, the CCD was chosen over the BBD because of it sequential nature and ease of interpretation.

Collection of the data for the model took place over two days in the faculty lounge of the Engineering Building. The design consisted of a full factorial CCD with two centre points and face-centred axial points. Two replicates of each point were performed to account for human error when shooting the golfer, for a total of 171 runs.

Analysis

The initial results were analyzed, to identify any outliers. In order to do this, the suggested fit from Design-Expert™ was used to analyze the data. A square root transform was applied to the data and DFFITS and Cook's D plots were used to determine any out-lying points. Any data points that appeared to be outliers were replicated and checked. If the new point produced a better fit, then it was used to replace the original data point.

Once all the outliers were corrected, a more thorough analysis was conducted. First the fit summary was reviewed to determine which type of model was a best fit for the data. Design-Expert™ suggested that a 2FI model be employed; however, the quadratic model would produce a better lack of fit. The model was backward selected using a quadratic process order. The selection process removed all 2FI and quadratic terms except for the A^2, AC and AD interactions. That left the model with the following factors: A, B, C, D, AC, AD and A^2. The p-value for the A^2 term was greater than 0.05 and was therefore removed from the model. This left a 2FI model with backwards selection, which was the suggested model by Design-Expert™.

The results of the ANOVA of the distance data (with a square root transform) are shown in Table 2. The ANOVA shows that the model is significant with the most important factors being angle and arm length. There is also a strong interaction between the arm length and angle. The lack of fit and error terms are not significant and have a large number of degrees of freedom. These all indicate a more than adequate model for the DOE Golf Toy.

Source	Sum of Squares	df	Mean Square	F Value	p-value Prob > F
Model	6394	10	639	678	< 0.0001
A-Angle	5580	1	5580	5914	< 0.0001
B-Direction	53.7	1	53.7	56.9	< 0.0001
C-Length	504	2	252	267	< 0.0001
D-Weights	137	2	68.7	72.9	< 0.0001
AC	103	2	51.5	54.6	< 0.0001
AD	15.7	2	7.86	8.34	0.0004
Residual	151	160	0.94		
Lack of Fit	75.7	70	1.08	1.29	0.1260
Pure Error	75.3	90	0.84		
Cor Total	6545	170			

Table 2 - ANOVA for Distance Model

The interaction term between arm length and angle (illustrated in Figure 1) indicated that, at larger arm lengths, the distance is greater for the given angle. Also, the distance vs. angle plot is steeper, which would indicate a small error in setting the angle at higher arm lengths would produce a greater error in distance. For this reason, when selecting optimum settings for the DOE Golf Toy, the settings with the lowest arm length were selected if both solutions were equally desirable.

The resultant formula for the distance is given in Equation 1.

$$\text{Distance} = (16.47 + 7.19A - 0.71B - 2.19C[1] + 0.18C[2] - 1.06D[1] - 0.074D[2] - 1.23AC[1] + 0.074AC[2] - 0.52AD[1] + 0.13AD[2])^2 \quad (1)$$

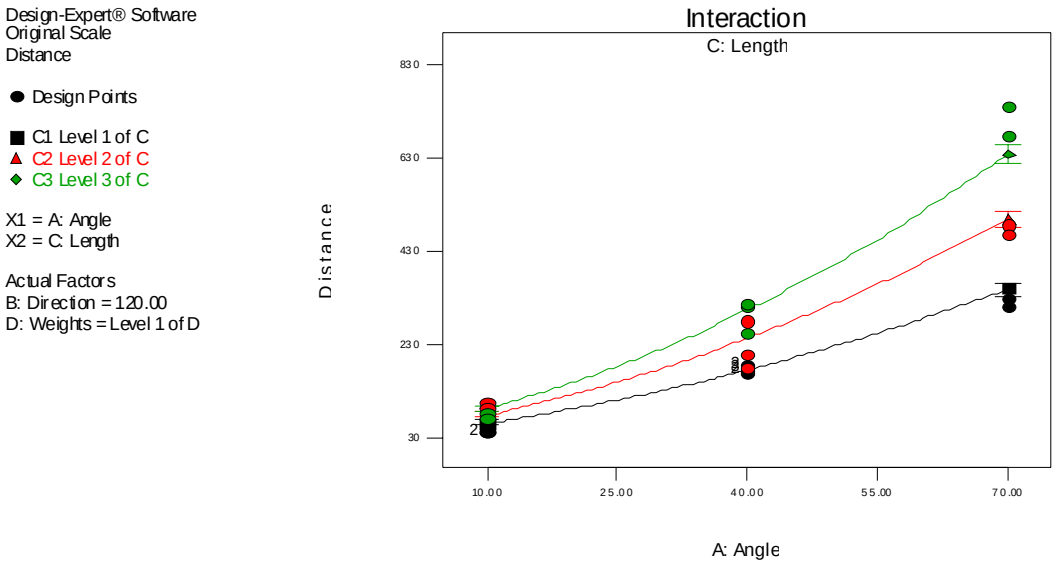

Figure 1 - Angle-Length (AC) Interaction

The actual vs. predicted plot, shown in Figure 2, indicates a good correlation between the selected model and the DOE Golf Toy, but there is considerable scatter in the values. This was noticed while the model was being tested; that, given the same setting, different distances could be achieved. This would indicate that there is some random error in the system that could not be modeled. This random error seemed to be independent from the distance, and would have to be compensated for during the golf tournament.

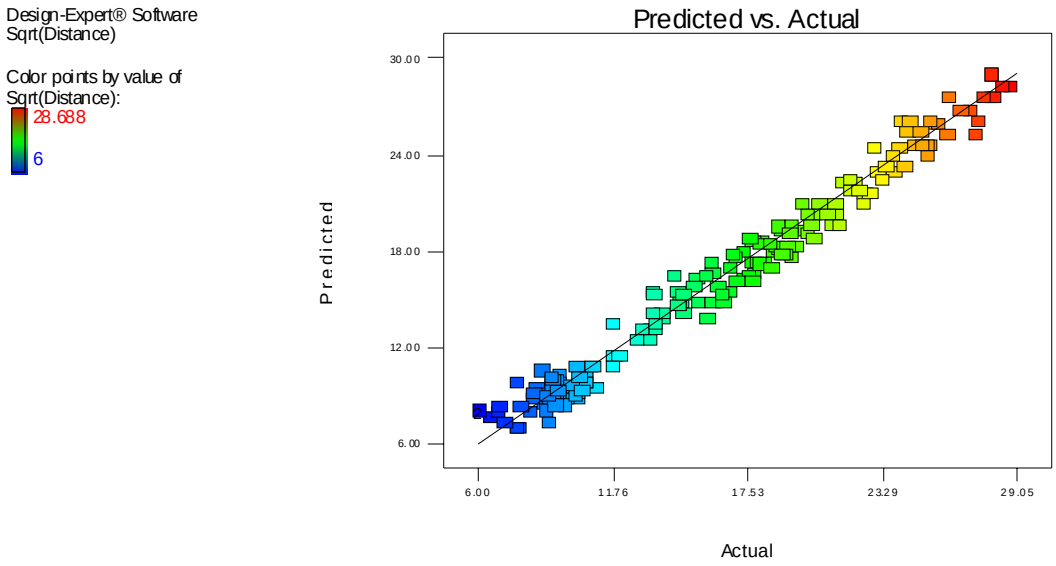

Figure 2 - Predicted vs Actual Distances

The deviation data was processed in a similar manner. There were no outliers to be removed from this data set because we replaced the deviation data when the outliers for distance were replaced. The ANOVA is shown in Table 3.

A $(y+154)^{2.11}$ power transform was applied to the deviation data and a model was backward-selected on a quadratic process order. The significant terms were found to be A, B, AB, C (based on hierarchy), BC, and B^2. The lack of fit and error terms were all insignificant and, therefore, the model was determined to be acceptable.

Source	Sum of Squares	df	Mean Square	F Value	p-value Prob > F
Model	15220127562	8	1902515945	22.24	< 0.0001
A-Angle	472534136	1	472534136	5.525	0.0200
B-Direction	9451438528	1	9451438528	110.5	< 0.0001
C-Length	4342581	2	2171291	0.025	0.9749
AB	1886717751	1	1886717751	22.06	< 0.0001
BC	1024305157	2	512152579	5.988	0.0031
B^2	2380789408	1	2380789408	27.84	< 0.0001
Residual	13855575237	162	85528242		
Lack of Fit	6585476555	72	91464952	1.132	0.2866
Pure Error	7270098682	90	80778874		
Cor Total	29075702800	170			

Table 3 - ANOVA for Deviation Model

Even though the model was found to be significant, the actual vs. predicted plot (shown in Figure 3) did not display a very good correlation. It was decided that, when selecting the actual deviation, the magnitude of the value should be considered, as well as the deviation calculated at other solution points. For example, if other solution points show a deviation of 15 cm, while the chosen solution has a deviation of 50 cm, then it should be considered that the chosen solution has over-predicted the deviation.

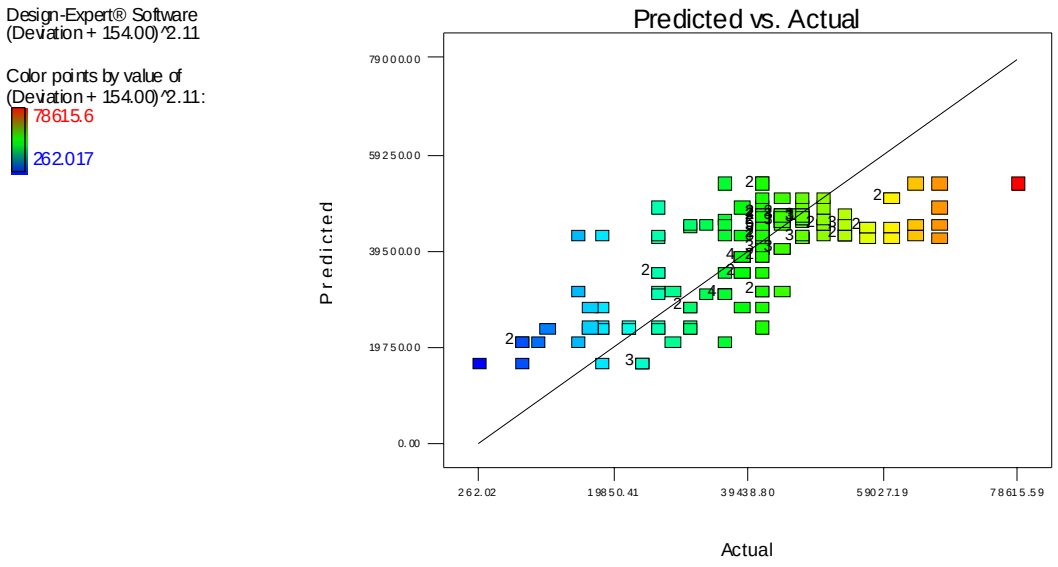

Figure 3 - Predicted vs Actual Deviation

Conclusion

The tournament illustrated how it is impossible to develop a perfect model, as there are uncontrollable factors that are not easily measured. However, the stability and robustness of the winning team's model showed that it is possible to achieve a reliable model if properly designed.

While conducting the experiment, it was concluded that for short distances the model would not answer as efficiently as for long distances. This spawned the development of a separate model specialized for short distances. However, it was observed that for very short distances it would be better to follow the experiences acquired during the practice rather than model, itself.

Regardless of how good a model is, without an expert and confident golfer no team can lead in a competition. Our exclusive shooting style based on the relaxation concept proved to be very essential in the final results, when compared to other groups. On this regard, yoga might help to improve the results.