Examination of Recovery Time Versus Average Speed for Finishers of the 2008 WEC

The average speeds inclusive of recovery time and exclusive of recovery time were treated as dependent variables; the independent variables were chooses as total recovery time, recovery time gates 1-6 inclusive, and gates 2-6, inclusive. The correlation coefficient between the sets of variables (2 x 3) was used to determine statistical strength for a linear fit. Nearly a linear fit was reported for the top 25 performers.

There were clear differences between the top 25 and the rest of the field. The correlation was strongest between gates 2-6, inclusive, and the average speed, including recovery time. The correlation coefficient was greater than -0.65 for the data set consisting of an array length of the first 25 finishers. The correlation coefficient was -0.73 for the top 10 finishers versus recovery time including gates 2-6. Since the final recovery time after finish met the FEI criterion < 30 minutes to pulse of 64 for all finishers, the a prior estimate was that the total recovery time, including final vet check would not be a strongly correlated as gate times 1-6 inclusive and gates 2-6 inclusive. Evaluating the correlation coefficients substantiated this argument. Another a priori assumption was that gates 2-6 inclusive would be more strongly correlated with average speed than gates 1-6 inclusive. This was also substantiated. One might attribute this to nervousness of the equine at the first pulse gate.

Furthermore, there was a clear distinction between the top 15 horses and the rest of the field. For the top 10 horses, as the horse recovered more quickly, the average speed including recovery time as well as the average speed excluding recovery time was significantly great with correlation coefficient approaching 0.7 for both cases. Pulse recovery was 64 for this event.

Using a correlation coefficient with a value of unity as the figure of merit, the (2x3) correlation matrix was as follows:

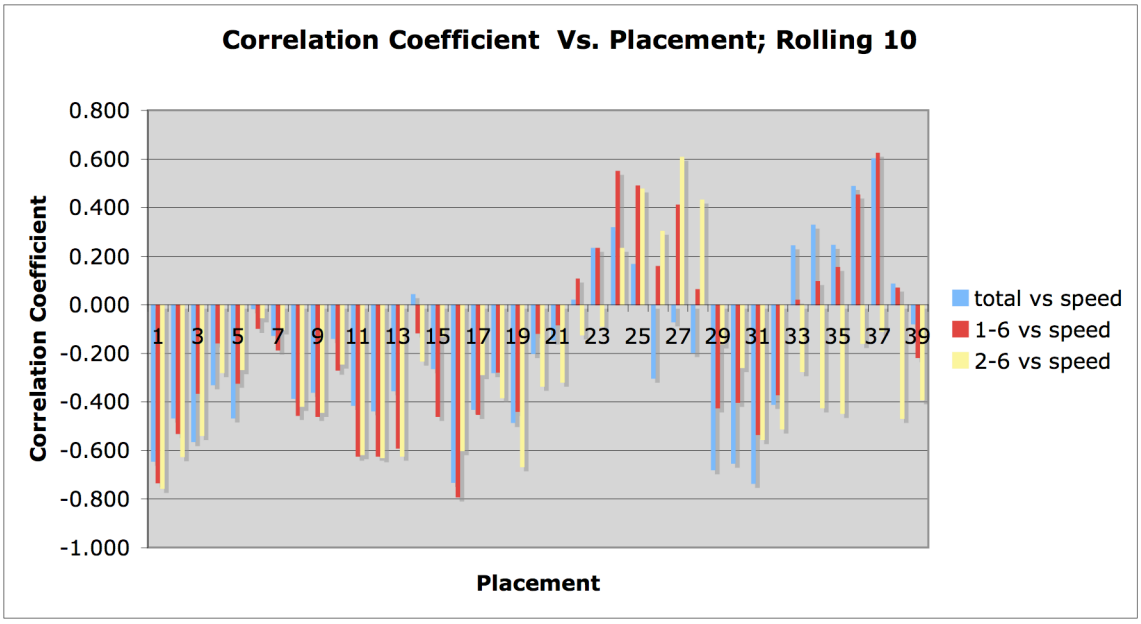
	1:F	1:06	2:06	
avg	-0.64	-0.60	-0.62	all 48
avg-recovery	-0.53	-0.48	-0.52	
avg	-0.64	-0.70	-0.73	top 25
avg-recovery	-0.37	-0.44	-0.50	
avg	-0.64	-0.73	-0.76	top 10
avg-recovery	-0.30	-0.39	-0.43	
avg	-0.08	0.06	-0.19	bottom 25
avg-recovery	0.16	0.28	0.13	
avg	-0.08	-0.22	-0.39	bottom 10
avg-recovery	0.38	0.28	0.09	

The variables of interest were independent (recovery time total, gates 1-6 inclusive, and gates 2-6, inclusive) and dependent (average speed inclusive and average speed exclusive of recovery time). In all cases of negative correlation; i.e., the average speed

was faster for faster recovery, the correlation was stronger relative to average speed inclusive of recovery time. There was a clear dichotomy between the top 10, the top 25, and the bottom 25; the average speed was much more highly correlated with the recovery time total. Also, for all cases of negative correlation, as expected, the correlation was stronger for recovery times inclusive of gates 2-6 rather than 1-6 or total recovery.

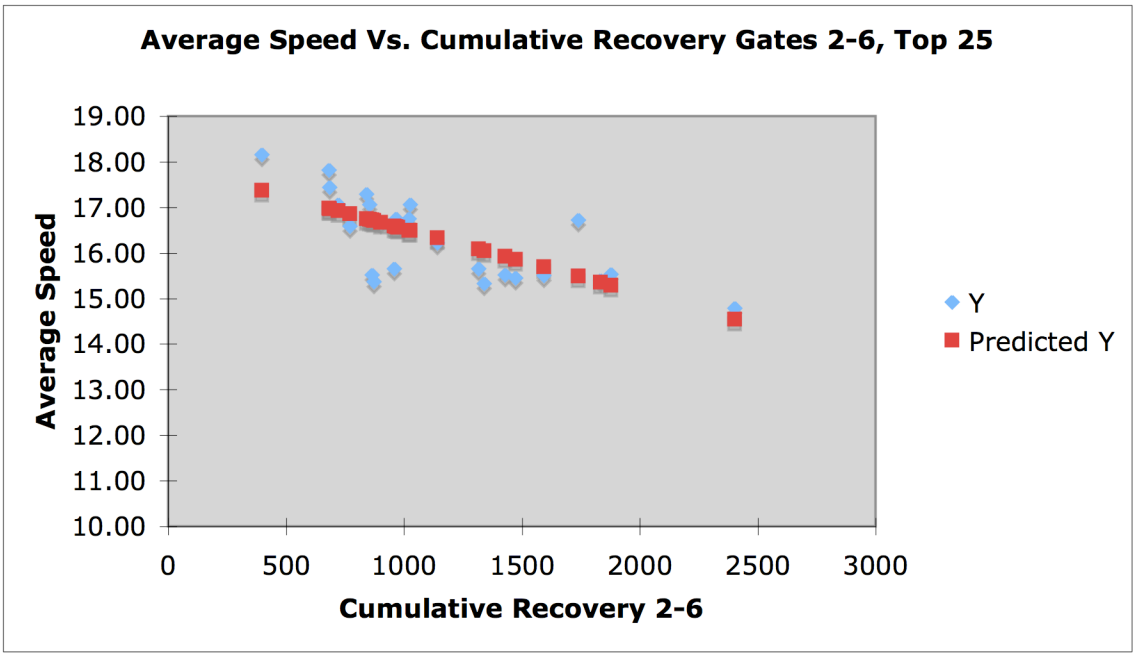
For the rest of the analyses, only a rolling average of 10 was examined for the top 25 finishers. The top 25 finishers showed a much stronger correlation of average speed relative to faster recovery; this correlation, if there, was much weaker for the bottom 25 finishers.

Using the more strongly correlated variables of recovery inclusive from gates 2-6 and average speed, inclusive of recovery time, the following chart displays the rolling x 10 correlation coefficient:



Since the rolling x 10 correlation coefficient approaches zero for a leading index of 22, the first 25 were clearly more able to handle speed at pace and recover quickly.

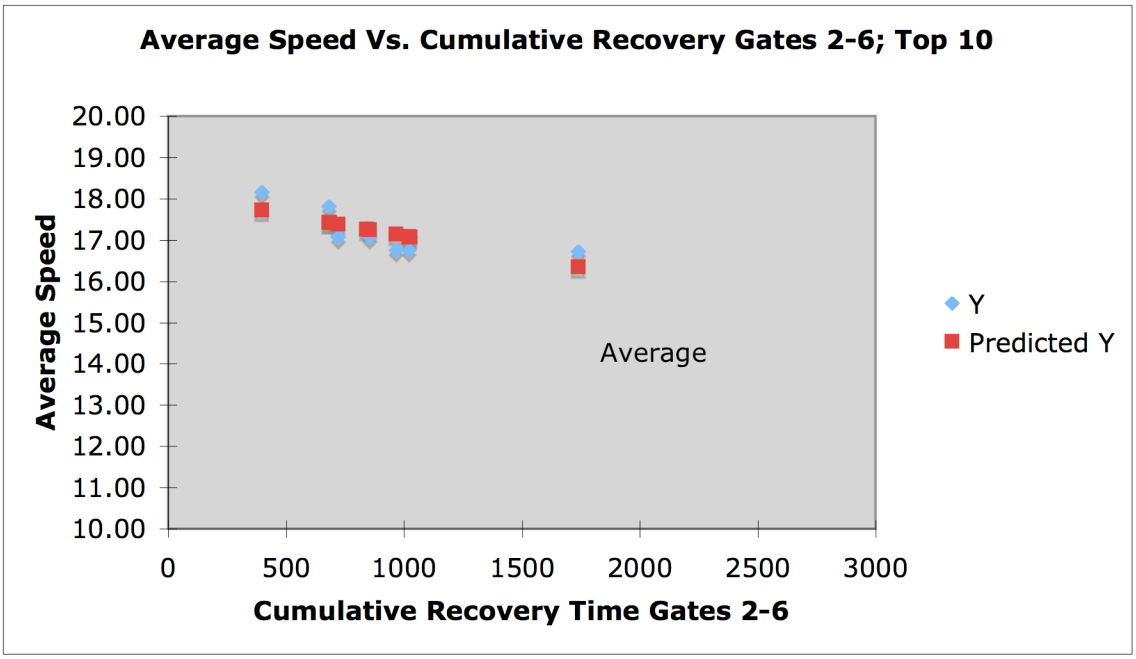
It is instructive to perform a regression analysis on the most highly correlated data; average speed versus recovery gates 2-6 for both the top 25 and the top 10. Considering the former,



ANOVA						
	df	*SS*	*MS*	*F*	*Significance F*	
Regression	1	10.37781142	10.37781142	25.6901608	3.93593E-05	
Residual	23	9.291092585	0.403960547			
Total	24	19.668904				
	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	17.9393597	0.343337792	52.24988375	2.08949E-25	17.22911137	18.64960803
X Variable 1	-0.001409676	0.000278122	-5.068546222	3.93593E-05	-0.001985015	-0.000834336

The X intercept shows the highest average speed one might predict from the recovery data and regression analysis applied to the top 25 finishers. The expected value is 17.9 km/hour with a 0-P standard deviation of 0.34 km/hour. The actual fastest average speed predicted within the standard deviation interval is some 18.3 km/hour. The actual highest actual speed was 18.16 km/hour.

A similar analysis was performed for the top 10 finishers.



ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	1.178539686	1.178539686	10.69078946	0.011361693
Residual	8	0.881910314	0.110238789		
Total	9	2.06045			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	18.13043234	0.29901612	60.63362841	6.08291E-12	17.44089993	18.81996475
X Variable 1	-0.00102512	0.000313523	-3.269677272	0.011361693	-0.001748106	-0.000302134

A way one might apply this to a particular competitor is to look at the deviation from average or just to use the slope predicted either from the top 10 or the top 25.

Let' examine the 24th finisher for example. This finisher was measured with a recovery time of 1340 seconds summed for gates 2-6. Using these regression values from the above tables, one would predict an average speed of 16.0 km/hour based on the top 25 or an average speed of 16.8 km/hour based on the top 10. The actual average speed was 15.3 km/hour, which was significantly influenced by a near 30-minute recovery at gate 1. Using the smaller value, a finishing time of 10.0 hours would be predicted. This competitor would then be expected to finish either 14th or 15th, a substantial change in placement.

Applying the same analysis to the 25th finisher with a measured 2-6 recovery time of 2400 seconds, the regression analysis indicates a time between 11.0 hours and 10.2 hours. In essence, this simple statistical analysis is predictive but not complete.

What is interesting the prediction the top few equines ran very close to the predicted maximum average speed based on these statistics. For training and selection purposes, recovery time is a strong predictor of race performance.

Patrick Perkins
2/28/09