Multivariate Modeling Of Gas Production Using Geologic and Completion Factors

Richard R. Batsell (Jones Graduate School, Rice University)
Sanjay Paranji (Anadarko)
Jason Mintz (Formerly of Anadarko, Now of Apache)

Prepared for Applied Geoscience Conference
March 6-8, 2018
Houston Geological Society

RICE | BUSINESS

Multivariate Modelling Of Gas Production Using Geology And Completion Factors

Richard R. Batsell (Jones Graduate School, Rice University), Sanjay Paranji (Anadarko),
Jason Mintz (Formerly of Anadarko, Now of Apache)

Important Managerial Questions

Does target zone B, which costs $200,000 a well more to complete than target zone A, produce enough additional gas to justify the extra cost? Does production increase linearly with increases in lateral length? Does proppant A produce more gas than proppant B? Holding constant geologic and completion factors, does fracking company A produce better results than fracking company B? Which is better: Sliding Sleeve or Plug and Perf? With enough cumulative experience across producing wells and sufficiently accurate models of production, the above questions, and many others, can be addressed.

Answering The Questions

Using data from 270 gas wells in the Marcellus Formation of Pennsylvania, this presentation illustrates a methodology for developing multivariate models of production. First, using all 40 available geology variables, 7 underlying key factors were extracted. These 7 geologic factors explained 50% of the variance in 180 day cumulative production. Then 20 completion variables were factor analyzed yielding 7 underlying completion factors. When the 7 completion factors were combined with the 7 geologic factors, the explained variance for 180 Day Cum increased to 64%.

Each of the 14 variables in the model are associated with a t statistic which reflects the: 1) direction of the effect of the variable on production; 2) whether or not the variable is statistically significant; and, if so, 3) the relative impact of the variable. One can thus assess the contribution of each variable to production and, eventually whether money invested in that variable yields production commensurate with the investment.

Finally, with such strong goodness-of-fit results, in essence simultaneously controlling for geology and completion factors, one can test questions like those in the first paragraph. So, target zone B's production was not significantly better than target zone A; production did increase linearly with lateral length; there were some fracking approaches that lead to significantly better results; and, Plug and Perf outperformed Sliding Sleeve.

This presentation will show the methodology as applied to these 270 wells and describe these and other important tests of managerial questions.

Figure 1
Frequency Distribution Of 90 Day Cumulative Gas (Mcf) Production Across 270 Wells

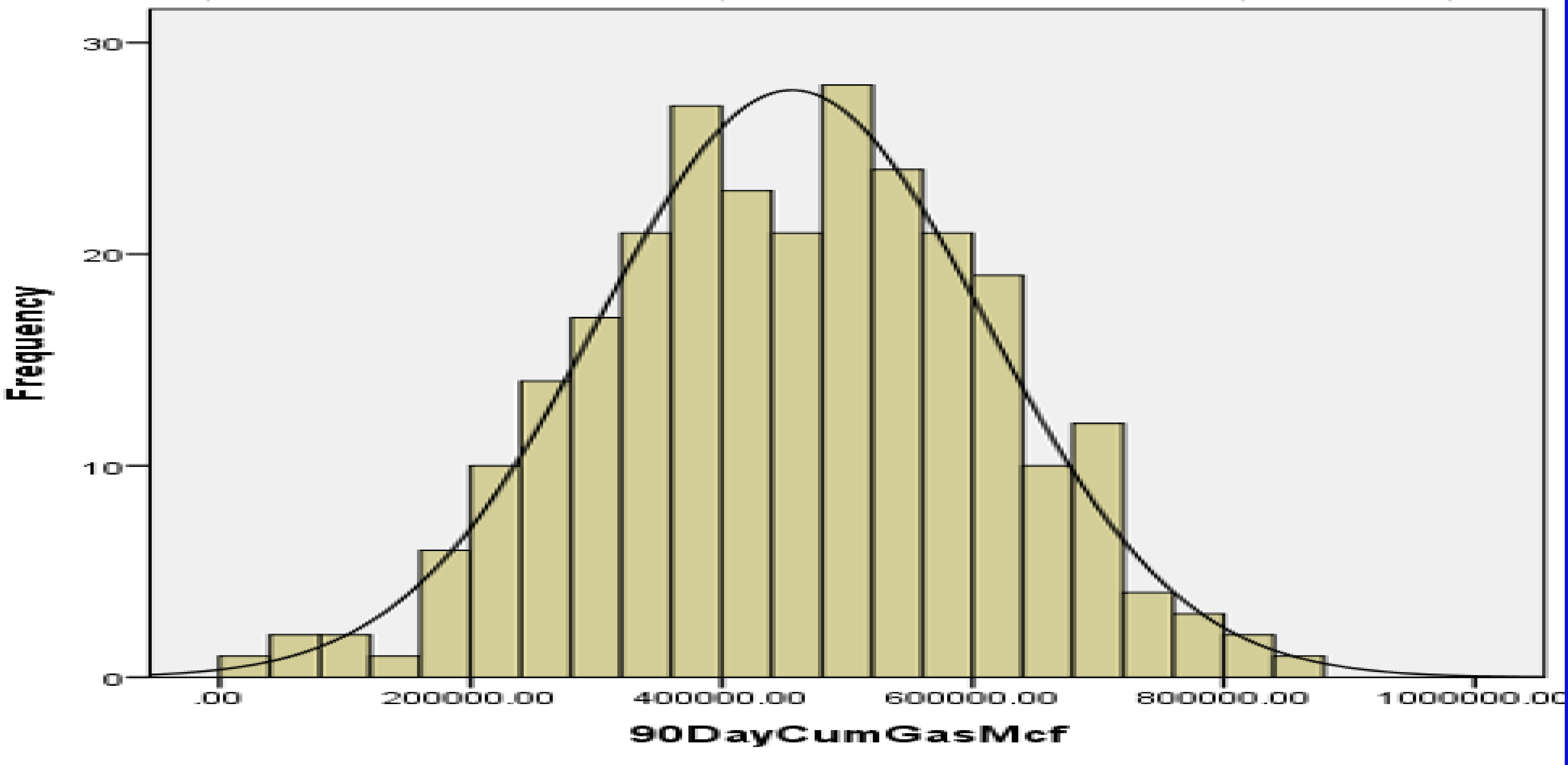


Figure 1 shows that the 90 day Cumulative Gas Production is very close to being normally distributed

Table 1
The 40 Geologic Variables Used In The Analysis

Variables				
HeelTVD		LWRMRCLAVGCORETOC		LWRMRCLEXCLUBERAVGVCLAY
ToeTVD		LWRMRCLAVGTOCWTP		LWRMRCLEXCLUBERAVGVSAND
AverageTVD		LWRMRCLGROSSISOPACH		LWRMRCLEXCLUBERGROSSISO
MRCLAVGCORETOC		LWRMRCLAVGRHOB		UBERMRCLAVGCOREPERM
MRCLAVGVITRINITEREFLECTANCE		LWRMRCLAVGVP		UBERMRCLAVGPHIG
FaciesBCSCycleThickness		LWRMRCLAVGVP12000FPS		UBERMRCLPHIGH
FaciesNETBCS		LWRMRCLAVGDEEPRESIS		UBERMRCLAVGPR
LWRMRCLAVGCOREPERM		LWRMRCLVLIME		UBERMRCLAVGSW
LWRMRCLAVGPHIG		LWRMRCLAVGPHIT		UBERMRCLAVGVCLAY
LWRMRCLPHIGH		LWRMRCLEXCLUBERAVGCOREPE		UBERMRCLAVGVSAND
LWRMRCLAVGPOISSONRATIO		LWRMRCLEXCLUBERAVGPHIG		UBERMRCLGROSSISOPACH
LWRMRCLAVGSW		LWRMRCLEXCLUBERPHIGH		UBERMRCLAVGCORETOC
LWRMRCLAVGVCLAY		LWRMRCLEXCLUBERAVGPR		
LWRMRCLAVGVSAND		LWRMRCLEXCLUBERAVGSW		

Figure 2

The First Example Of Two Geology Variables Inversely Correlated (R = - . 835)

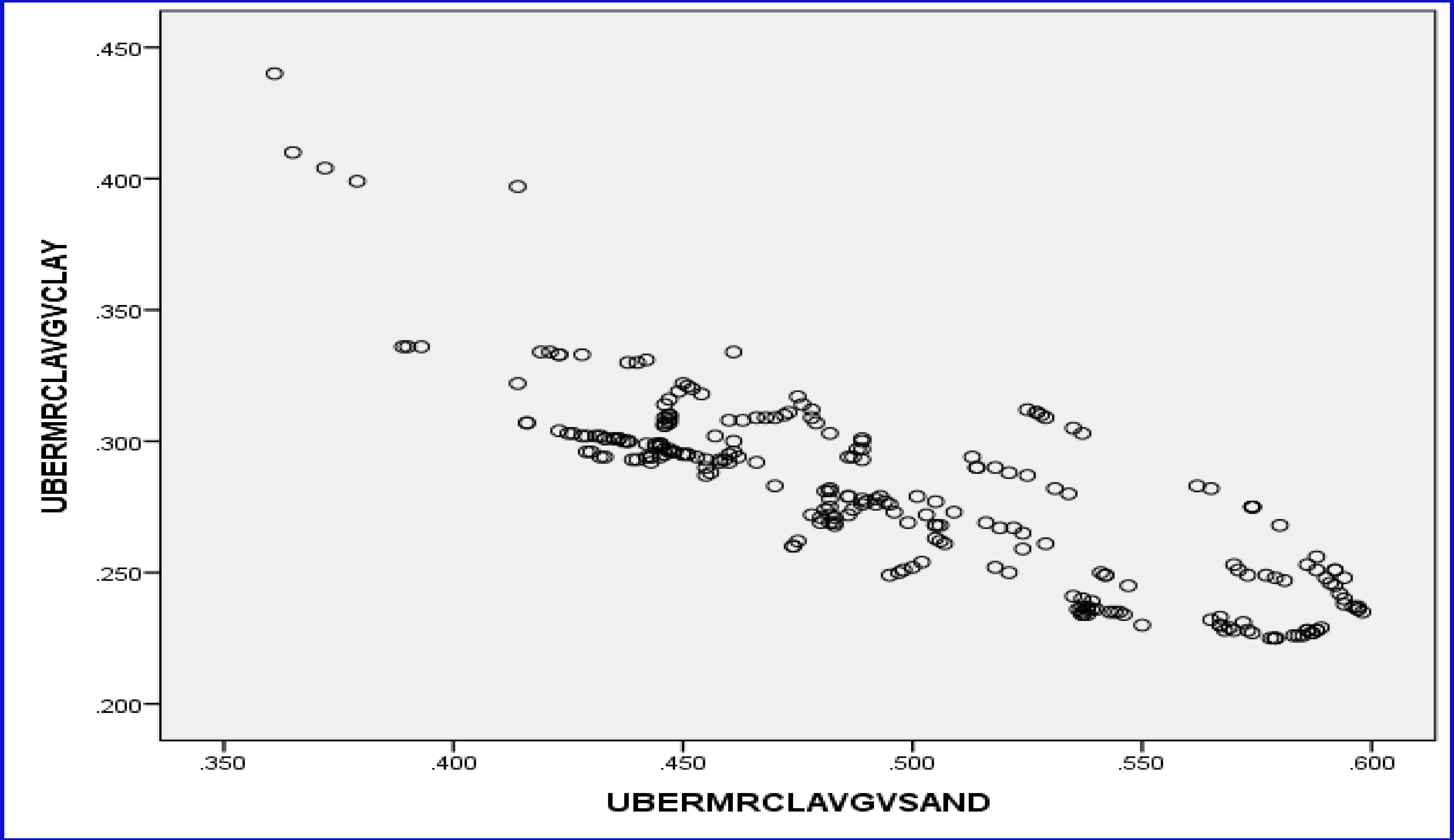
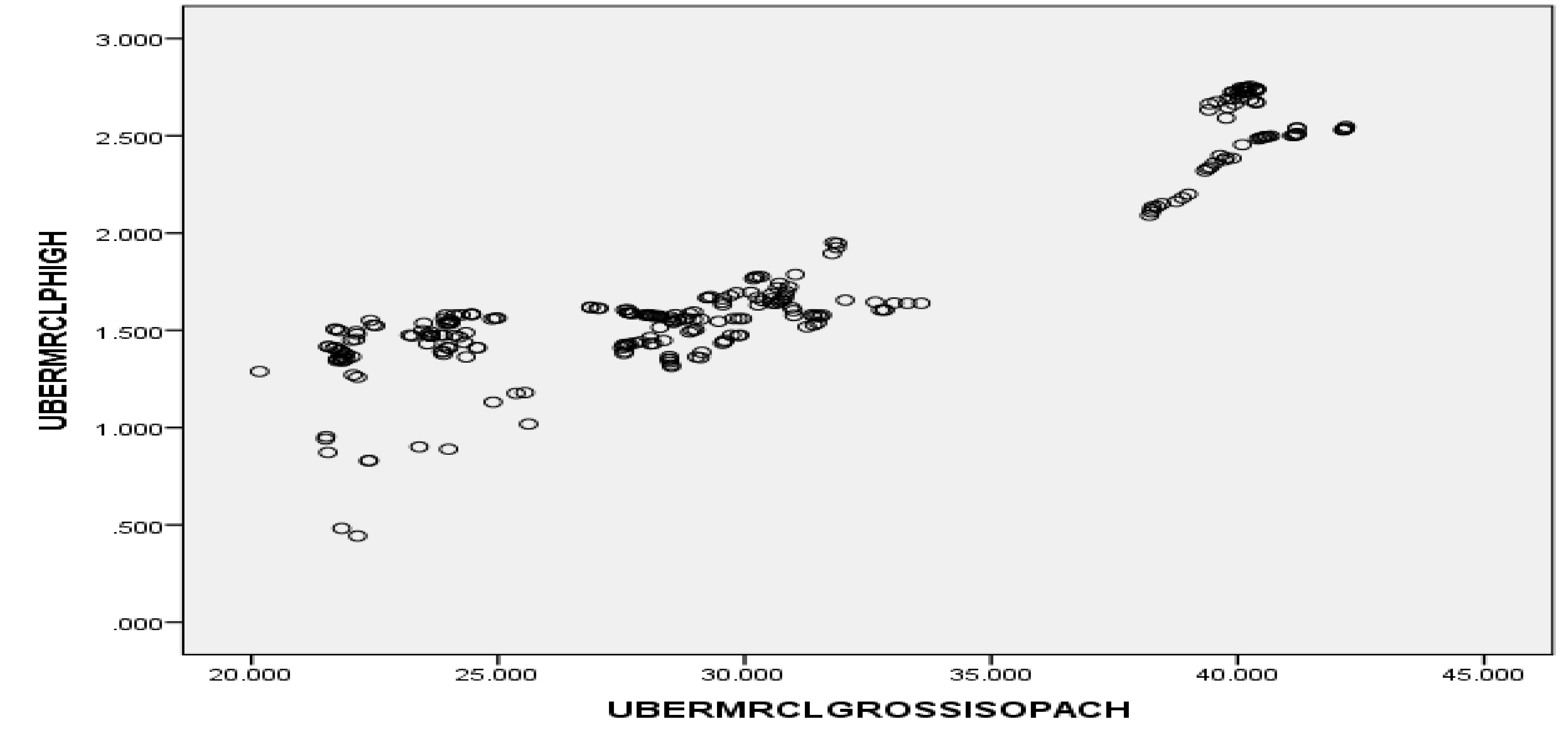


Figure 3

An Example of Two Geology Variables Positively Correlated (R = . 906)



Running simple multivariate models with independent variables this highly correlated, can produce unstable and difficult-to-interpret results.

Table 2
The Percent Of The Variance In The Original 40 Geologic Variables Explained By Each Of The Derived 7 Factors

Factors:	Percent Variance
GF1	23.611
GF2	20.911
GF3 TVD	14.731
GF4	9.261
GF5	7.619
GF6 Perm	7.331
GF7	6.294

Principle Components Factor Analysis identified 7 underlying dimensions present in the 40 original geology variables. Those 7 geology dimensions were rendered as 7 orthogonal new variables capturing 89.8% of the variance from the original 40.

Table 3

Regression Analysis Using the 7 Geologic Factor Scores to Fit the 30 Day, 60 Day, 90 Day, 180 Day, and 365 Day Production Numbers For Gas (Number of Observations, Model Fits, t statistics: Factors are ordered by importance for 365 Day)

	Cumulative Gas Mcf				
	30 Day	60 Day	90 Day	180 Day	365 Day
N=	270	270	269	264	211
Adj R^2=	42.8%	48.0%	48.9%	49.7%	50.8%
Variables:					
GF1	9.927	11.2390	11.5670	12.1730	10.1200
GF6Perm	8.034	9.050	9.056	8.290	5.327
GF3TVD	4.415	4.126	4.238	4.886	4.645
GF7	2.649	2.309	2.620	3.305	3.911
GF2	-	-	-	-2.169	-2.585
GF5	4.130	4.776	4.667	4.100	2.276
GF4	-	-	-	-	2.017

As can be seen in Table 3, the resulting multivariate model explains almost 50% of the variance in 180 Day cumulative production (Adjusted R^2 = 49.7%). The t statistics in this table can be used to assess the direction and relative impact of each geology factor.

Table 4
The 20 Completion Variables Used In The Analysis

Variables	
NetPerfdLateralLength	ClustersPerStage
CleanSlickWaterVolbbls	NumberofPerfsperCluster
ActualTotalFluidbbls	AvgInjRateAllStages
ActualProppantlbs	CleanSlickWaterperLateralFoot
NumberofStages	ActualTotalFluidperLateralFoot
TotalPerfs	ActualProppantperLateralFoot
LateralLengthperNumberStages	ActualAcidperLateralFoot
GrossLateralLengthperStage	ActualAcidVolbbls
Total100Mesh	AvgPerfsperStage
ClusterSpacing	Tortuosity

Table 5
The Percent Of The Variance In The Original 20 Completion Variables Explained By Each Of The Derived Completion Factors

Factors:	Percent Variance
CF1 LatLenVol	21.966
CF2 LenPerStage	17.400
CF3	16.921
CF4	16.629
CF5	9.281
CF6	5.899
CF7	5.223

The derived 7 completion factors explained 93.3% of the variance in the original 20 completion variables.

Table 6

Step 2: Regression Analysis Using the 7 Geologic and 7 Completion Factor Scores to Fit the 30 Day, 60 Day, 90 Day, 180 Day, and 365 Day Production Numbers For Gas (Number of Observations, Model Fits, t statistics: Factors are ordered by importance for 365 Day)

	Cumulative Gas Mcf				
	30 Day	60 Day	90 Day	180 Day	365 Day
N=	270	270	269	264	211
Adj R^2=	47.2%	55.4%	58.7%	63.8%	66.1%
Variables:					
GF1	10.2480	11.3960	12.2630	13.7610	10.1520
CF1LatLenVol	2.855	5.287	6.619	8.522	8.227
CF2LenPerStage	-2.918	-3.833	-4.235	-5.374	-5.344
GF3TVD	3.760	3.619	4.285	5.172	5.304
GF7	3.384	3.179	3.400	4.496	5.222
GF6Perm	7.788	8.927	9.206	9.006	5.147
GF2	-	-2.417	-3.270	-4.642	-4.393
GF4	2.392	2.868	2.819	3.447	3.827
GF5	4.625	5.299	5.128	4.733	2.678
CF4	-	-	-	-	-2.050
CF3	-2.869	-2.062	-	-	-
CF5	-	-	-	-	-
CF6	-	-	-	-	-
CF7	-	-	-	-	-

One Caution, One Generality, and 4 Additional Examples Of The Successful Application Of The Methodology In This Presentation

1) Care in defining your variables
2) Applies to oil as well as gas production
3) Marcellus versus Uber
4) Linearity of production
5) Plug and Perf versus Sliding Sleeve
6) Fracking Company A versus Fracking Company B