respect › commitment › teamwork

Updated Simulation of Unidentified Gas levels

DESC 16th Feb 2016

Introduction

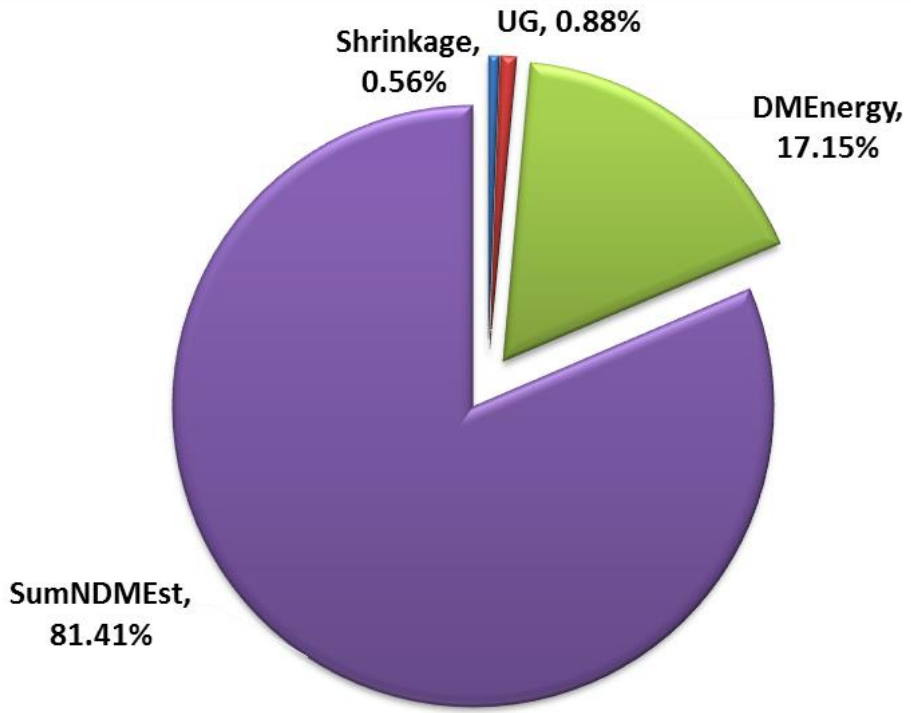
- There is a new concept of daily Unidentified Gas (UG) in the post Nexus regime
- We last presented our analysis on UG to DESC on 19th May 2015
- For background info on UG please refer to the previous presentation
- This presentation analyses updated UG results using the most recent data

Data used in the analysis

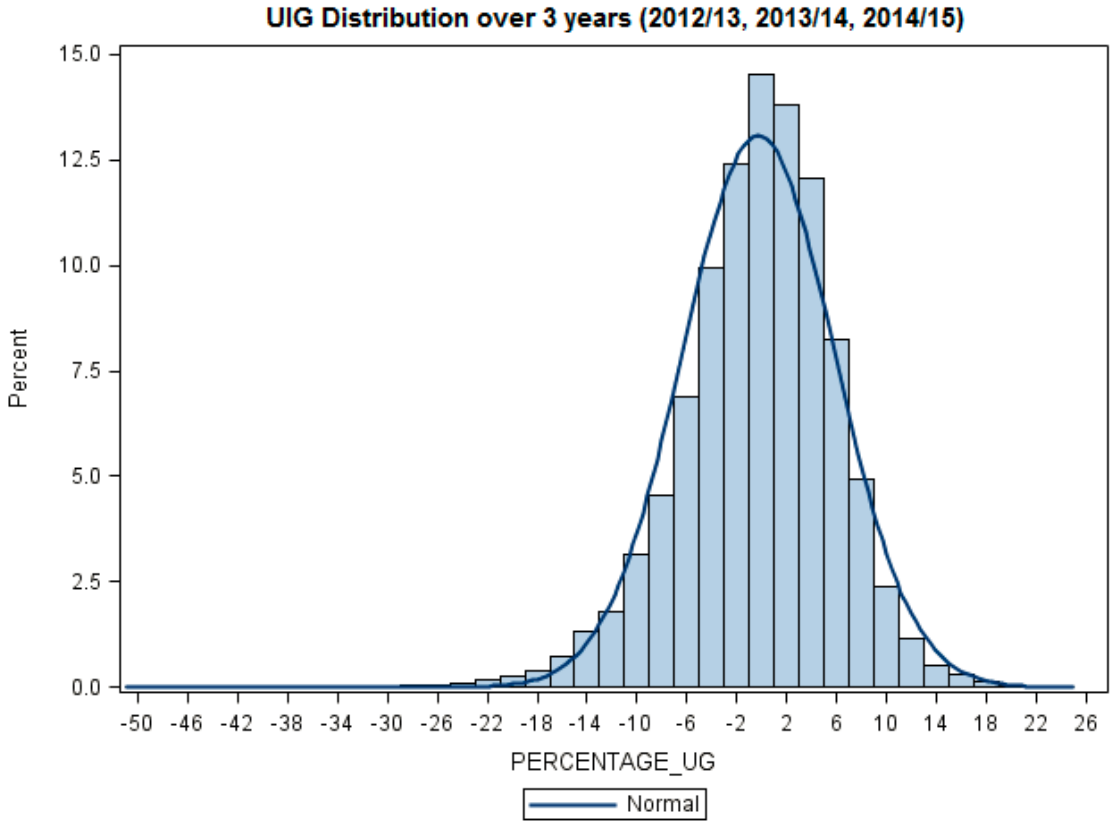
- Daily data at LDZ level
- Actual DM and LDZ measurements
- Actual NDM AQs
- Gas Years analysed:
  - 2012/13
  - 2013/14
  - 2014/15
- Simulated the new algorithm to derive estimates of what UG would have been using
  - Revised ALPs and DAFs under the new seasonal normal basis
  - Revised CWV definitions

Updated UG Proportion Summary

Total proportion (nationally) over 3 gas years



Updated National Distribution of UG



Mean = -0.29
Std Dev = 6.10
95% of the UG values are between **-12.49 and 11.91** (values outside of this range are often regarded as unusual).

We are **95%** confident that the population mean for UG is contained in the interval -0.39 and -0.19.
We are **90%** confident that the population mean for UG is contained in the interval -0.37 and -0.20

UG Analysis – Summary Statistics

UG Summary Statistics (%)				
	Mean	**Std Dev**	**Min**	**Max**
2012/13	1.47	5.93	-50.54	23.62
2013/14	-2.03	5.78	-33.95	19.93
2014/15	-0.31	6.08	-50.00	21.91

	Mean	**Std Dev**	**Min**	**Max**
2012/13, 2013/14, 2014/15	-0.29	6.10	-50.54	23.62

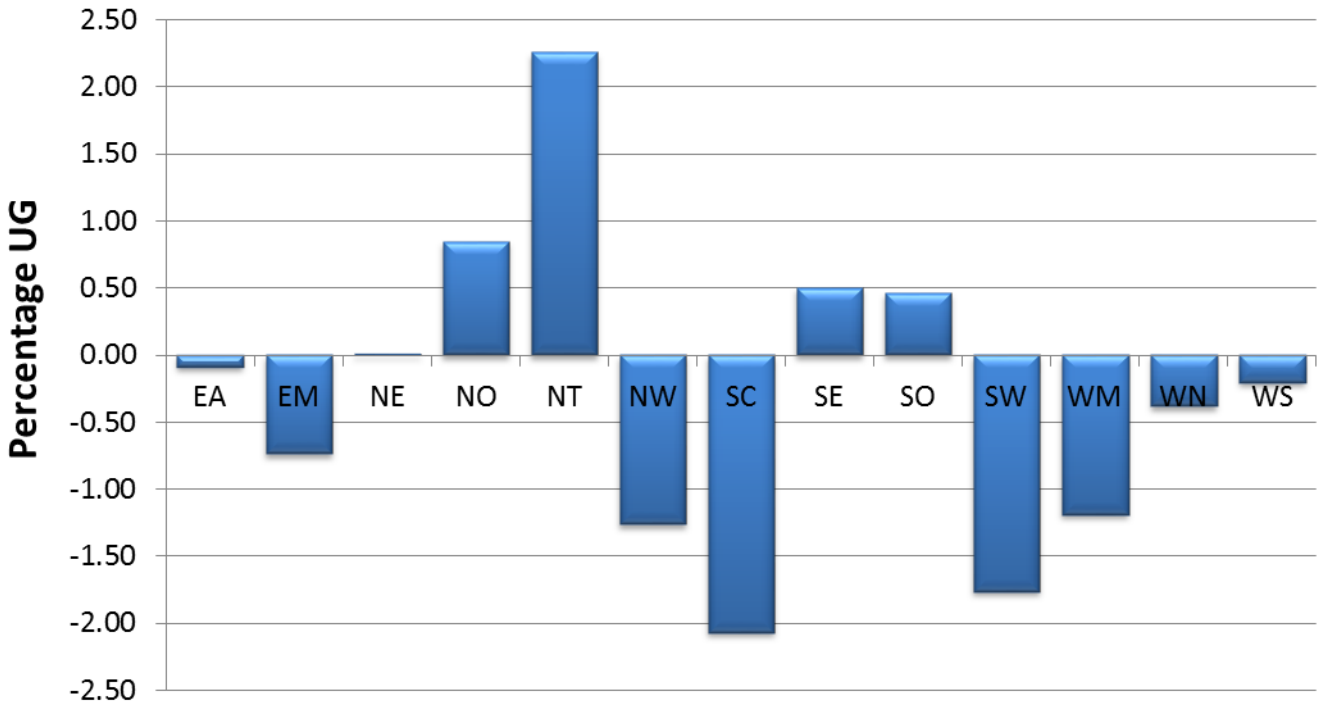
UG Analysis – Summary Statistics

UG Summary Statistics by LDZ				
LDZ	**Mean**	**Std Dev**	**Min**	**Max**
EA	-0.09	5.18	-17.46	16.61
EM	-0.74	5.19	-16.94	18.92
NE	0.01	5.65	-26.13	16.13
NO	0.85	5.44	-20.26	23.44
NT	2.26	5.38	-17.99	20.51
NW	-1.26	6.98	-29.51	16.85
SC	-2.07	6.13	-21.57	13.99
SE	0.51	5.64	-50.00	21.91
SO	0.46	5.42	-19.84	17.71
SW	-1.76	6.55	-26.02	18.16
WM	-1.19	5.91	-26.24	15.64
WN	-0.38	7.69	-27.74	19.34
WS	-0.21	6.17	-50.54	23.62

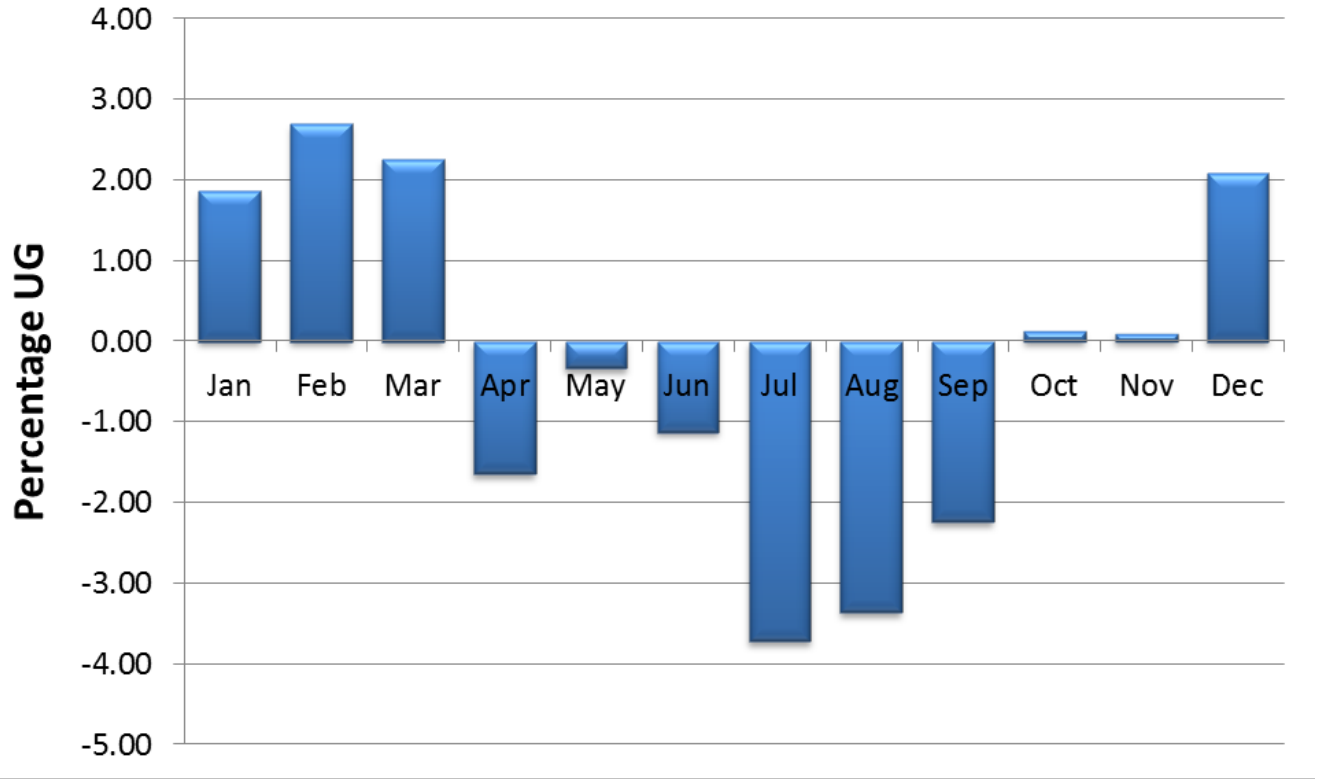
UG Summary Statistics by Month				
Month	**Mean**	**Std Dev**	**Min**	**Max**
Jan	1.86	3.67	-9.13	19.93
Feb	2.70	3.54	-8.24	13.61
Mar	2.26	5.79	-21.38	19.17
Apr	-1.64	8.43	-29.51	23.62
May	-0.33	7.15	-50.54	23.44
Jun	-1.12	6.22	-22.26	21.91
Jul	-3.71	5.47	-24.24	18.92
Aug	-3.36	5.53	-23.69	13.76
Sep	-2.23	7.26	-28.02	20.51
Oct	0.13	5.38	-50.00	16.05
Nov	0.10	4.08	-28.19	18.73
Dec	2.08	3.72	-33.95	14.40

The above tables display the summary statistics across the 3 years.

Average UG over last 3 years by LDZ



Average UG over last 3 years by month



GB CWV comparison over the 3 years

Month	Average GB CWV over 3 gas years	Average GB CWV 2012/13	Average GB CWV 2013/14	Average GB CWV 2014/15
Oct	11.35	9.99	12.01	12.05
Nov	6.95	6.53	6.33	7.98
Dec	4.65	4.01	5.37	4.56
Jan	3.72	3.02	4.61	3.52
Feb	3.86	2.94	4.98	3.67
Mar	5.78	3.48	7.43	6.43
Apr	9.87	8.44	10.88	10.29
May	12.47	11.96	13.17	12.29
Jun	14.69	14.53	15.06	14.48
Jul	15.48	15.64	15.59	15.21
Aug	15.25	15.52	14.96	15.28
Sep	14.13	14.03	14.73	13.64

The cells highlighted in blue show the coldest month out of the 3 years.
The cells highlighted red show the warmest month out of the 3 years.
On average, it appears the 2012/13 was the coldest year and 2013/14 was the warmest year.

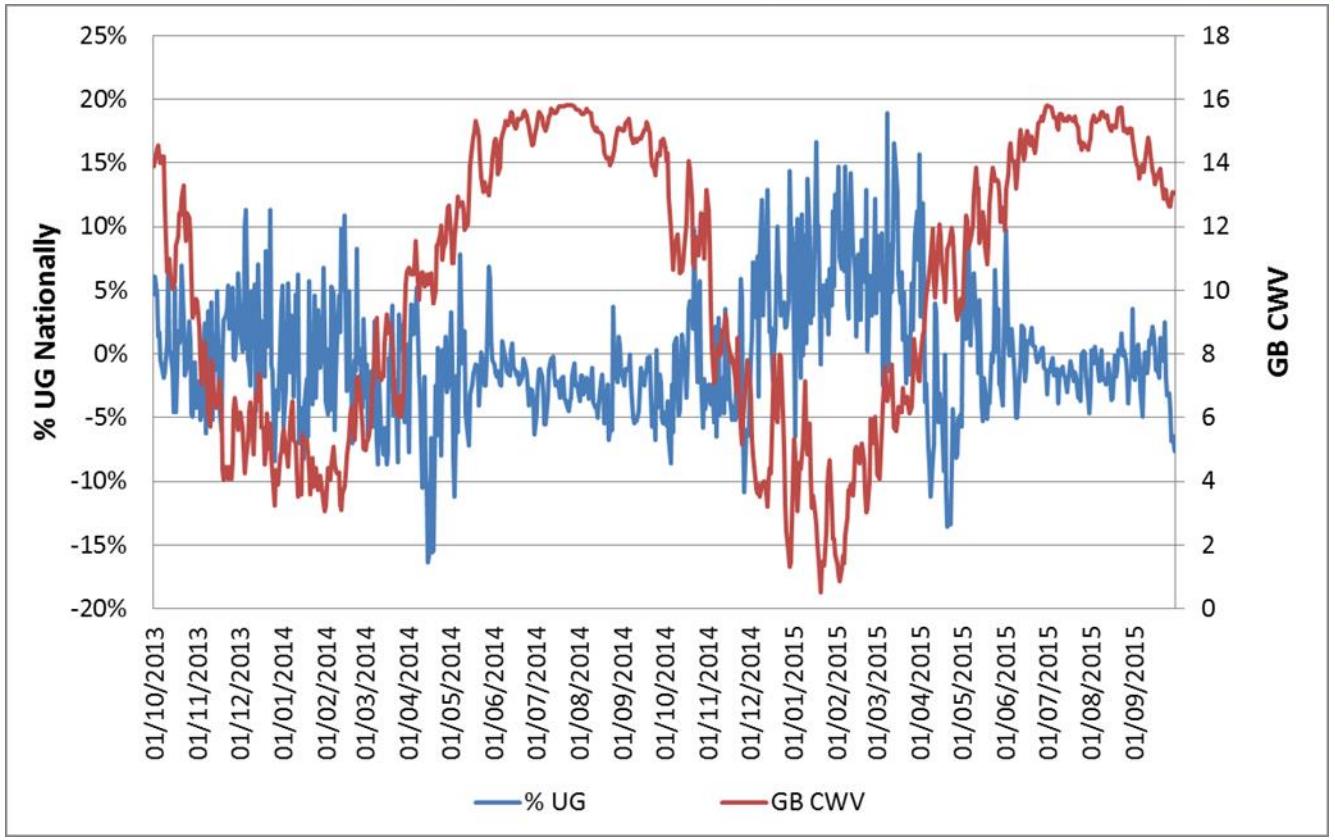
Simple Linear Regression results

Simple linear regression was carried out on the following variables to investigate if the variation in UG could be explained.
The data used in the analysis was at LDZ level over the 3 gas years

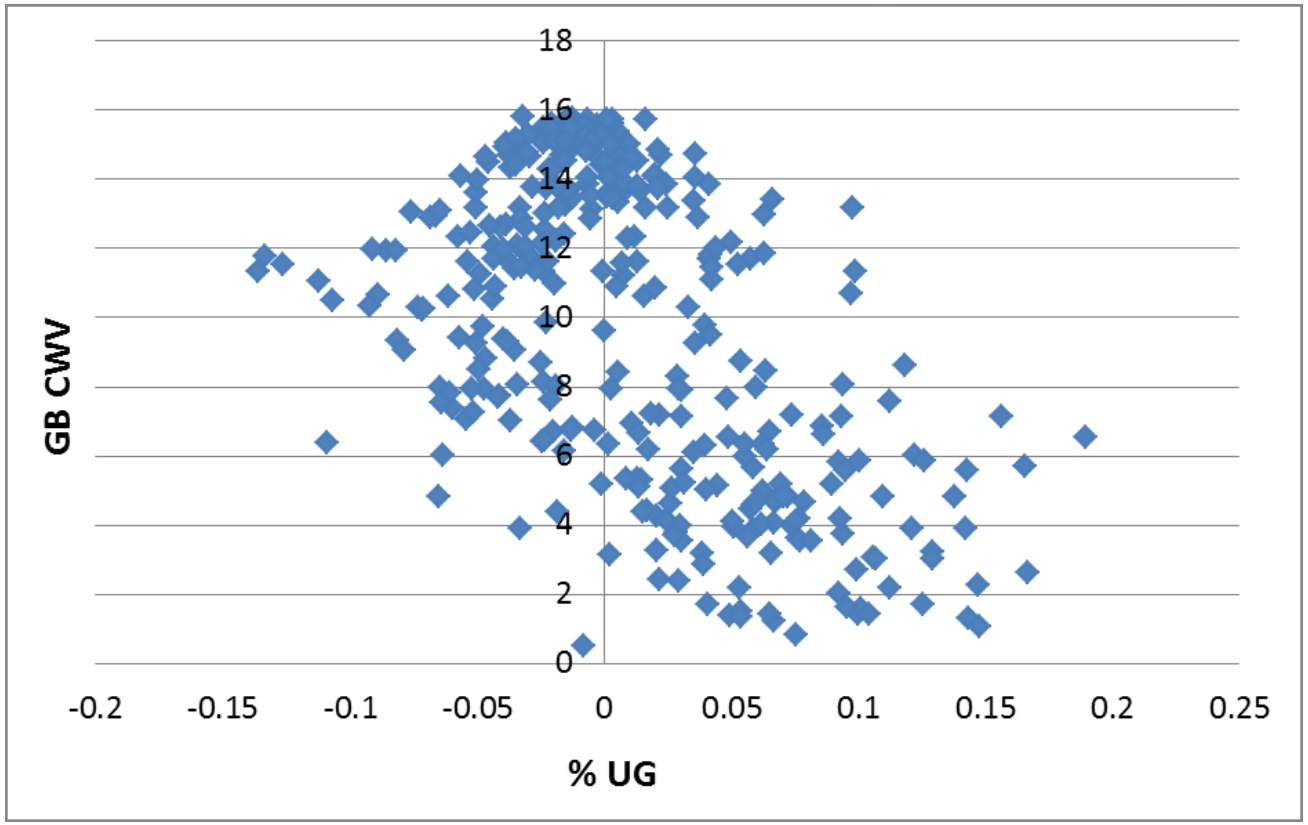
x	*y*	R^2
CWV	UG	12.97%
SumNDMEST	UG	11.44%
DMEnergy	UG	0.88%
SNCWV	UG	8.07%
WCF	UG	8.81%
EUC01B	UG	12.08%

The following graphs also allow us to explore any possible relationships. (note: GB CWV and UG values has been used in the charts for visual purposes only. The analysis was carried out at LDZ level).

Exploring the relationship between UG and CWV



Exploring the relationship between UG and CWV



Correlation analysis between GB CWV and % UG
Oct 2013 – Sep 2015 (latest 2 gas years)

xoserve

respect › commitment › teamwork

Multiple Regression Analysis

- After investigating the individual variables through simple linear regression, it was clear that there is large proportion of variation in unidentified gas that is still unexplained.

- Further analysis was then carried out which allowed for several independent variables to be investigated simultaneously.

  It also allowed for other variables to be explored (e.g. day of the week and month).

- The first model to be explored using multiple regression contained all of the variables tested in simple linear regression earlier, to see how well they performed together.

Multiple Regression Results

The REG Procedure
Model: MODEL1
Dependent Variable: UIG

Number of Observations Read	14235
Number of Observations Used	14235

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	6	1.286254E17	2.143757E16	500.22	<.0001
Error	14228	6.097586E17	4.285624E13		
Corrected Total	14234	7.38384E17			

Root MSE	6546468	R-Square	0.1742
Dependent Mean	1035583	Adj R-Sq	0.1739
Coeff Var	632.15281		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	1558855	298286	5.23	<.0001
Shrinkage	1	-1.94871	0.40641	-4.79	<.0001
DMEnergy	1	-0.00095400	0.00553	-0.17	0.8630
SumNDMEst	1	-0.03744	0.00830	-4.51	<.0001
cwv	1	-147321	25619	-5.75	<.0001
WCF	1	-867321	38432	-22.57	<.0001
EUC01B	1	0.08043	0.01088	7.39	<.0001

Model Selection

- The number of variables to be used in the analysis = 54
  (this includes all weather variables, demand variables at LDZ and EUC level, dummy variables for: day of the week, holidays and month)

- With such a large number of variables, eliminating one variable at a time using standard multiple regression can take an extreme amount of time.

- Due to the large number of variables, Stepwise Regression seems to be a sensible automated method to select the best model.

Stepwise Regression

- Stepwise Regression starts with an empty model and incrementally builds a model one variable at a time. Variables already in the model will not necessarily remain (like Forward selection). The Backward component of the method removes variables from the model that do not meet the significance criteria (0.05)

- When carrying out the analysis, the best model gave an R^2 of 27.83%. There is still a large proportion of variation in UG that is still unexplained.

Conclusions

- There is still a large proportion of variation in UG that is unexplained.
- There does not appear to be a strong relationship between CWV and UG.
- UG is most likely to be negative in the summer months.