Electronic Supplementary Material (ESI) for Environmental Science: Nano.

Supplementary Information

Table S1 Details of MeOx (hydroxide) NP dataset.

CompNo	D1_Metals	D2_SemiMetals	D3_HeteroNonMetals	Metals_SumIP	SemiMetals_SumIP	Electrons_ActiveM	AtWt_ActiveM	MassNo_ActiveM	Neutons_ActiveM	IP_ActivM	X_ActivM	VWR_ActivM	Electrons_Activ_SM
*1	1	0	2	577.4	0	13	26.982	27	14	577.4	1.5	0.143	0
2	1	0	2	577.4	0	13	26.982	27	14	577.4	1.5	0.143	0
3	1	0	2	577.4	0	13	26.982	27	14	577.4	1.5	0.143	0
4	1	0	2	577.4	0	13	26.982	27	14	577.4	1.5	0.143	0
*5	1	0	2	577.4	0	13	26.982	27	14	577.4	1.5	0.143	0
6	1	0	2	735	0	28	58.6934	59	31	735	1.8	0.124	0
7	1	0	2	735	0	28	58.6934	59	31	735	1.8	0.124	0
8	1	0	2	735	0	28	58.6934	59	31	735	1.8	0.124	0
*9	1	0	2	735	0	28	58.6934	59	31	735	1.8	0.124	0
10	1	0	2	735	0	28	58.6934	59	31	735	1.8	0.124	0
11	1	0	1	735	0	28	58.6934	59	31	735	1.8	0.124	0
12	1	0	1	904.5	0	30	65.38	65	35	904.5	1.6	0.138	0
*13	1	0	1	904.5	0	30	65.38	65	35	904.5	1.6	0.138	0
14	1	0	2	526.8	0	58	140.12	140	82	526.8	1.1	0.181	0
15	1	0	2	526.8	0	58	140.12	140	82	526.8	1.1	0.181	0
16	1	0	2	526.8	0	58	140.12	140	82	526.8	1.1	0.181	0
*17	1	0	2	526.8	0	58	140.12	140	82	526.8	1.1	0.181	0
18	1	0	2	526.8	0	58	140.12	140	82	526.8	1.1	0.181	0
19	1	0	2	526.8	0	58	140.12	140	82	526.8	1.1	0.181	0
20	1	0	2	658	0	22	47.867	48	26	658	1.5	0.147	0
*21	1	0	2	658	0	22	47.867	48	26	658	1.5	0.147	0
22	1	0	2	658	0	22	47.867	48	26	658	1.5	0.147	0
23	1	0	2	658	0	22	47.867	48	26	658	1.5	0.147	0
24	0	1	2	0	786.3	0	0	0	0	0	0	0	14
*25	0	1	2	0	786.3	0	0	0	0	0	0	0	14

To cont…

AtWt_Activ_SM	MassNo_Activ_SM	Neutons_Activ_SM	IP_Activ_SM	X_Activ_SM	VWR_Activ_SM	Electrons_Activ_NM	AtWt_Activ_NM	MassNo_Activ_NM	Neutons_Activ_NM	IP_Activ_NM	X_Activ_NM	VWR_Activ_NM	SuM_Active_M_SM
0	0	0	0	0	0	8	15.999	16	8	1314	3.5	0.074	1
0	0	0	0	0	0	8	15.999	16	8	1314	3.5	0.074	1
0	0	0	0	0	0	8	15.999	16	8	1314	3.5	0.074	1
0	0	0	0	0	0	8	15.999	16	8	1314	3.5	0.074	1
0	0	0	0	0	0	8	15.999	16	8	1314	3.5	0.074	1
0	0	0	0	0	0	8	15.999	16	8	1314	3.5	0.074	1
0	0	0	0	0	0	8	15.999	16	8	1314	3.5	0.074	1
0	0	0	0	0	0	8	15.999	16	8	1314	3.5	0.074	1
0	0	0	0	0	0	8	15.999	16	8	1314	3.5	0.074	1
0	0	0	0	0	0	8	15.999	16	8	1314	3.5	0.074	1
0	0	0	0	0	0	8	15.999	16	8	1314	3.5	0.074	1
0	0	0	0	0	0	8	15.999	16	8	1314	3.5	0.074	1
0	0	0	0	0	0	8	15.999	16	8	1314	3.5	0.074	1
0	0	0	0	0	0	8	15.999	16	8	1314	3.5	0.074	1
0	0	0	0	0	0	8	15.999	16	8	1314	3.5	0.074	1
0	0	0	0	0	0	8	15.999	16	8	1314	3.5	0.074	1
0	0	0	0	0	0	8	15.999	16	8	1314	3.5	0.074	1
0	0	0	0	0	0	8	15.999	16	8	1314	3.5	0.074	1
0	0	0	0	0	0	8	15.999	16	8	1314	3.5	0.074	1
0	0	0	0	0	0	8	15.999	16	8	1314	3.5	0.074	1
0	0	0	0	0	0	8	15.999	16	8	1314	3.5	0.074	1
0	0	0	0	0	0	8	15.999	16	8	1314	3.5	0.074	1
0	0	0	0	0	0	8	15.999	16	8	1314	3.5	0.074	1
28.085	28	14	786.3	1.8	0.132	8	15.999	16	8	1314	3.5	0.074	1
28.085	28	14	786.3	1.8	0.132	8	15.999	16	8	1314	3.5	0.074	1

To cont…

SuMElectrons_Active_M_SM	SuMAtWt_Active_M_SM	SuMMassNo_Active_M_SM	SuMNeutons_Active_M_SM	cat_rad	hydration rate (x_h2o)	oxidation degree of metal (n_oxy_M)	zeta potential	solubility	SSA(surface area)	pauling electronegativity	hydroxide	oxide	LDH
13	26.982	27	14	67.5	0.5	3	37	-6.5	121	1.61	1	0	1.1
13	26.982	27	14	67.5	0.5	3	39	-6.5	107	1.61	1	0	1.08
13	26.982	27	14	67.5	0.5	3	37	-6.5	62	1.61	1	0	1.05
13	26.982	27	14	67.5	0.5	3	33	-6.5	40	1.61	1	0	1.07
13	26.982	27	14	67.5	0.5	3	37	-6.5	54	1.61	1	0	1.07
28	58.6934	59	31	83	1	2	40	-3.5	190	1.91	1	0	2.76
28	58.6934	59	31	83	1	2	44	-3.5	226	1.91	1	0	2.57
28	58.6934	59	31	83	1	2	31	-3.5	192	1.91	1	0	2.99
28	58.6934	59	31	83	1	2	31	-3.5	86	1.91	1	0	2.2
28	58.6934	59	31	83	1	2	35	-3.5	10	1.91	1	0	2.07
28	58.6934	59	31	83	0	2	35	-12	40	1.91	0	1	0.88
30	65.38	65	35	88	0	2	15	-2.5	93	1.65	0	1	2.62
30	65.38	65	35	88	0	2	29	-2.5	19	1.65	0	1	2.81
58	140.12	140	82	101	0	4	-26	-9.5	151	1.12	0	1	2.02
58	140.12	140	82	101	0	4	-30	-9.5	52	1.12	0	1	2.58

58	140.12	140	82	101	0	4	-35	-9.5	61	1.12	0	1	2.08
58	140.12	140	82	101	0	4	-17	-9.5	60	1.12	0	1	0.98
58	140.12	140	82	101	0	4	-31	-9.5	37	1.12	0	1	2.08
58	140.12	140	82	101	0	4	9	-9.5	61	1.12	0	1	0.82
22	47.867	48	26	74.5	0	4	9	-9	94	1.54	0	1	0.88
22	47.867	48	26	74.5	0	4	1	-9	302	1.54	0	1	1.01
22	47.867	48	26	74.5	0	4	-4	-9	45	1.54	0	1	1.29
22	47.867	48	26	74.5	0	4	0	-9	44	1.54	0	1	1
14	28.085	28	14	54	0	4	-9	-2.7	326	1.9	0	1	1.2
14	28.085	28	14	54	0	4	-29	-2.7	447	1.9	0	1	1.09

*Presents test set compounds Me_{Ox} NPs

Additional Statistical Results

1. **Test for normality of residuals and goodness of fit.**

   Kolmogorov –Smirnov test and **Shapiro-Wilk** were performed to check the normality distribution of residuals. The results obtained from the test run in SPSS show that p values are > 0.05 suggesting acceptance of H_0 i.e., residuals are normally distributed.

   M1

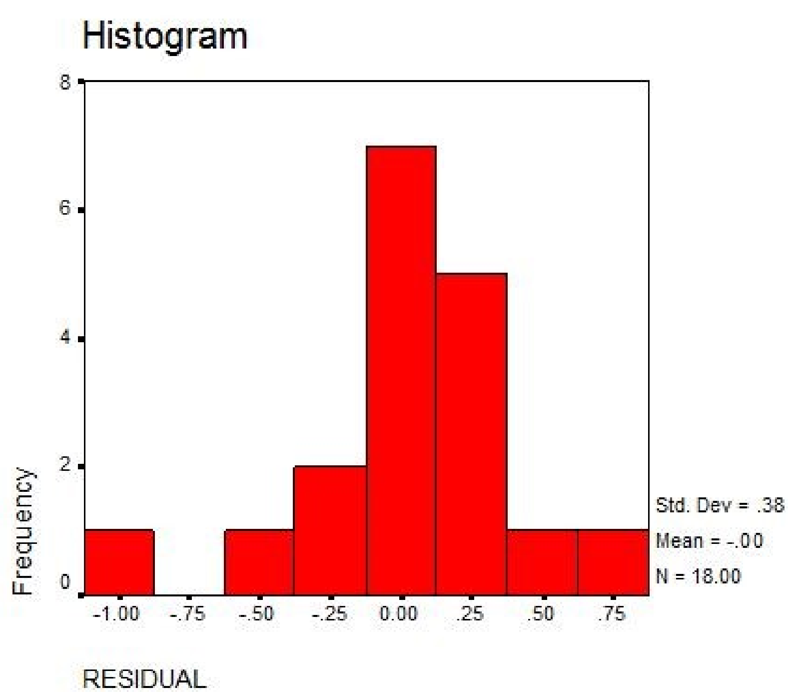
Tests of Normality

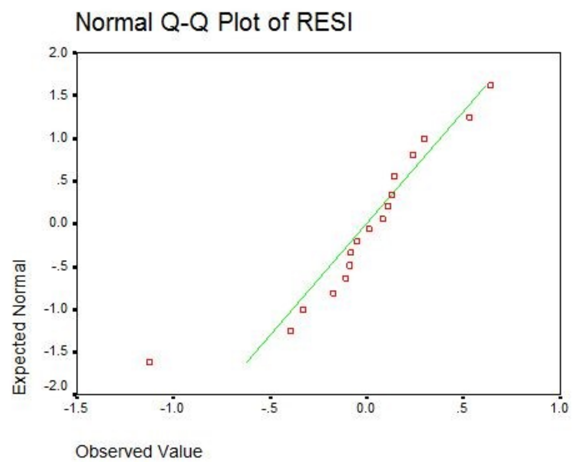
	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
RESIDUAL	.166	18	.200*	.901	18	.062

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

RESIDUAL





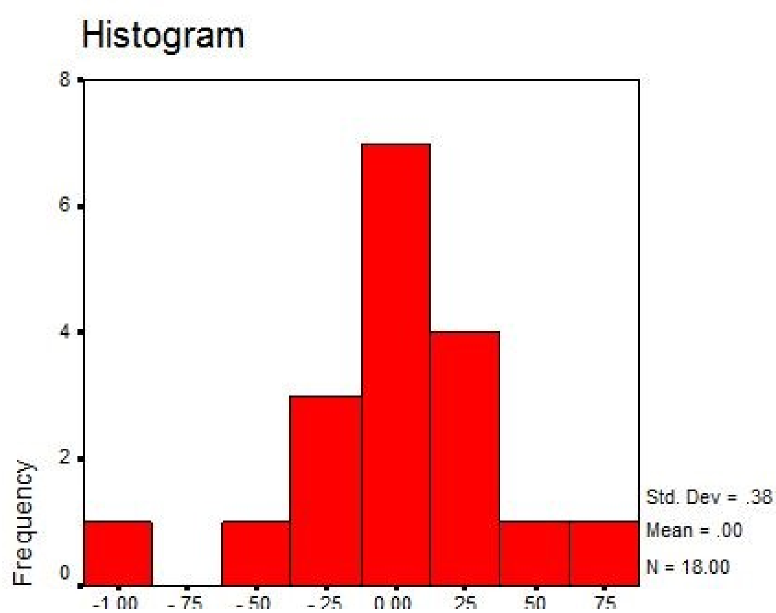
M2

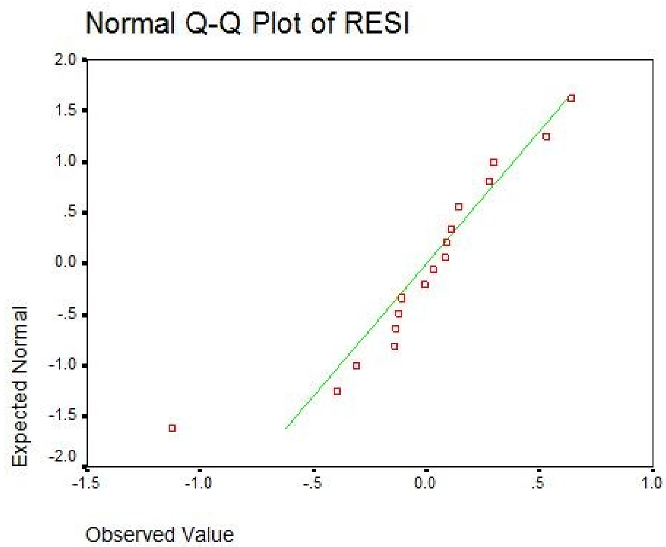
Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
RESIDUAL	.190	18	.085	.901	18	.061

a. Lilliefors Significance Correction

RESIDUAL





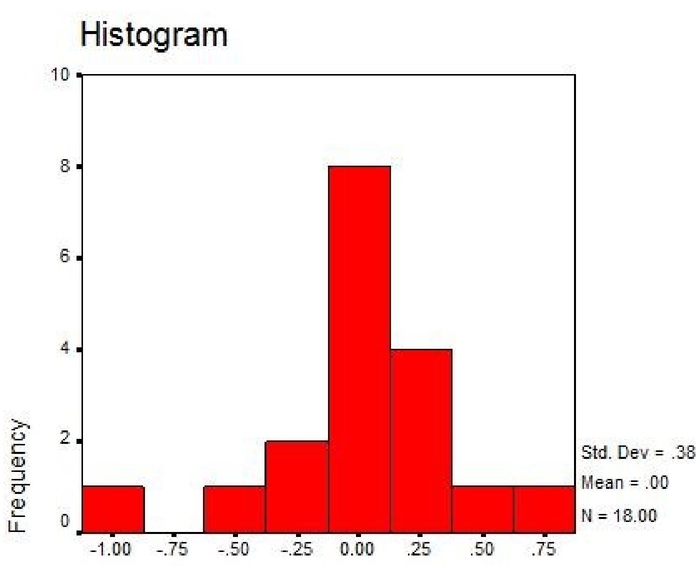
M3

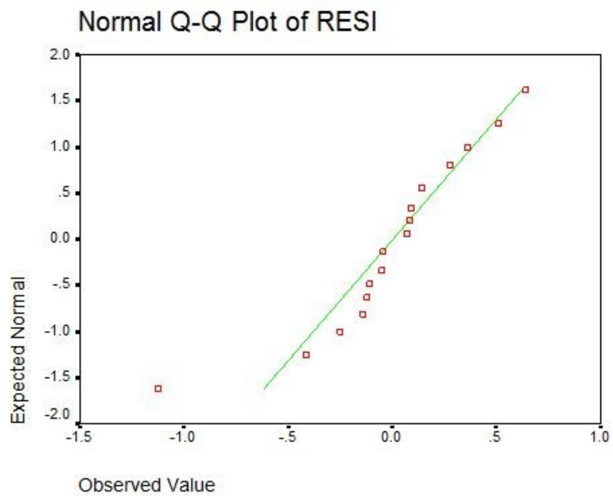
Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
RESI	.188	18	.091	.898	18	.054

a. Lilliefors Significance Correction

RESI





M4

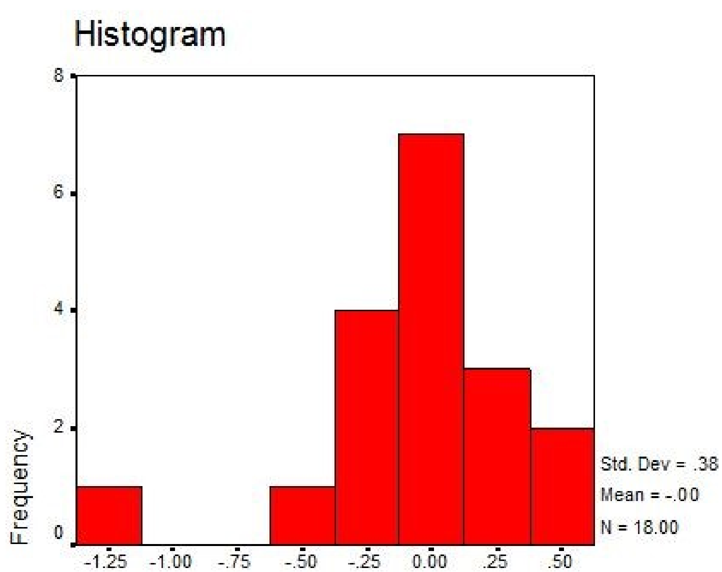
Tests of Normality

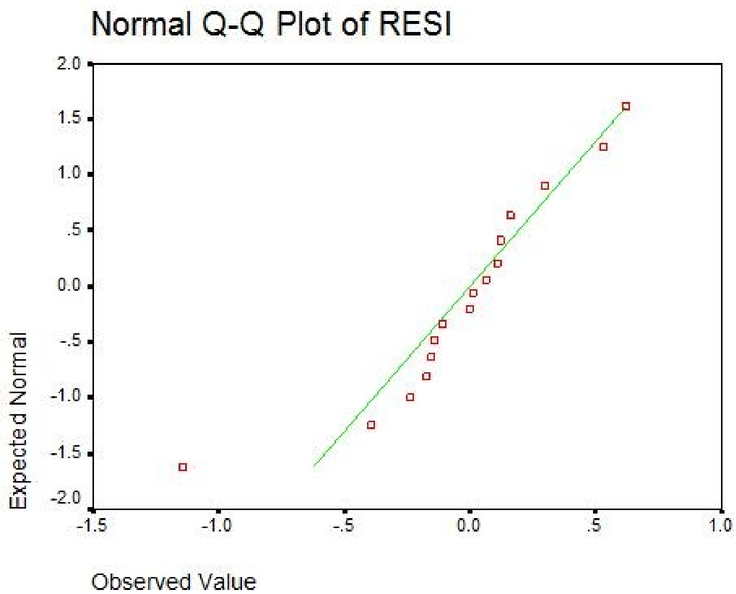
	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
RESI	.163	18	.200*	.895	18	.047

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

RESI





2. Non multicollinearity

Further co-linearity was also checked among the descriptors obtained from the four models with the help of SPSS software. The VIF (variable inflation factor) is checked here. The value of VIF starts with 1 and has no upper limit. The value 1 indicates there is no correlation between a given predictor variable and any other predictor variable in the model. We see none VIF value is greater than 5, so multicollinearity will not be a problem here.

M1

Coefficients[a]

Model		Unstandardized Coefficients B	Std. Error	Standardized Coefficients Beta	t	Sig.	Collinearity Statistics Tolerance	VIF
1	(Constant)	1.256	.307		4.094	.001		
	VAR00002	1.394E-03	.000	.392	3.598	.002	.996	1.005
	VAR00003	2.876E-02	.005	.653	5.321	.000	.783	1.277
	VAR00004	.197	.031	.788	6.436	.000	.786	1.272

a. Dependent Variable: VAR00005

Collinearity Diagnostics[a]

Model	Dimension	Eigenvalue	Condition Index	Variance Proportions (Constant)	VAR00002	VAR00003	VAR00004
1	1	3.665	1.000	.00	.01	.01	.01
	2	.185	4.453	.04	.22	.43	.06
	3	.105	5.917	.00	.08	.56	.74
	4	4.498E-02	9.027	.96	.70	.00	.19

a. Dependent Variable: VAR00005

M2

Coefficients[a]

Model		Unstandardized Coefficients B	Std. Error	Standardized Coefficients Beta	t	Sig.	Collinearity Statistics Tolerance	VIF
1	(Constant)	-.324	.463		-.699	.492		
	VAR00002	5.311E-04	.000	.149	1.235	.230	.809	1.237
	VAR00003	3.630E-02	.007	.703	5.318	.000	.674	1.484
	VAR00004	.193	.030	.772	6.373	.000	.804	1.243

a. Dependent Variable: VAR00005

Collinearity Diagnostics[a]

Model	Dimension	Eigenvalue	Condition Index	Variance Proportions (Constant)	VAR00002	VAR00003	VAR00004
1	1	3.797	1.000	.00	.01	.00	.01
	2	.139	5.221	.00	.20	.00	.61
	3	5.031E-02	8.687	.22	.71	.04	.28
	4	1.331E-02	16.891	.77	.08	.96	.10

a. Dependent Variable: VAR00005

M3

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.	Collinearity Statistics Tolerance	Collinearity Statistics VIF
1	(Constant)	.125	.577		.216	.831		
	VAR00002	3.815E-02	.006	.739	5.969	.000	.778	1.285
	VAR00003	-.129	.116	-.155	-1.116	.277	.619	1.615
	VAR00004	.174	.038	.695	4.575	.000	.517	1.934

a. Dependent Variable: VAR00005

Collinearity Diagnostics[a]

Model	Dimension	Eigenvalue	Condition Index	Variance Proportions (Constant)	VAR00002	VAR00003	VAR00004
1	1	3.843	1.000	.00	.00	.00	.01
	2	.100	6.189	.05	.03	.00	.53
	3	4.586E-02	9.155	.00	.12	.64	.18
	4	1.033E-02	19.293	.95	.84	.35	.29

a. Dependent Variable: VAR00005

M4

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.	Collinearity Statistics Tolerance	Collinearity Statistics VIF
1	(Constant)	-.281	.467		-.602	.554		
	VAR00002	.355	.368	.131	.966	.345	.655	1.527
	VAR00003	3.675E-02	.007	.712	5.203	.000	.646	1.547
	VAR00004	.207	.031	.826	6.685	.000	.793	1.260

a. Dependent Variable: VAR00005

Collinearity Diagnostics[a]

Model	Dimension	Eigenvalue	Condition Index	Variance Proportions (Constant)	VAR00002	VAR00003	VAR00004
1	1	3.841	1.000	.00	.00	.00	.01
	2	.104	6.086	.04	.02	.01	.94
	3	4.218E-02	9.543	.20	.83	.02	.02
	4	1.278E-02	17.335	.76	.14	.97	.03

a. Dependent Variable: VAR00005

3. No auto-correlation

The Durbin Watson (DW) statistic is a test to determine autocorrelation in SPSS software. The test assumes a value between 0 to 4 and the value DW=2 indicates that there is no autocorrelation. The DW value below 2 reflects that the data is positively autocorrelated. The results obtained in the present test shows clearly that the DW value is 2.091 which means that no autocorrelation in the data.

Regression

Variables Entered/Removed[b]

Model	Variables Entered	Variables Removed	Method
1	VAR00007, VAR00006, VAR00003, VAR00004, VAR00002, VAR00005[a]	.	Enter

a. All requested variables entered.

b. Dependent Variable: VAR00001

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.876[a]	.767	.689	.4176	2.091

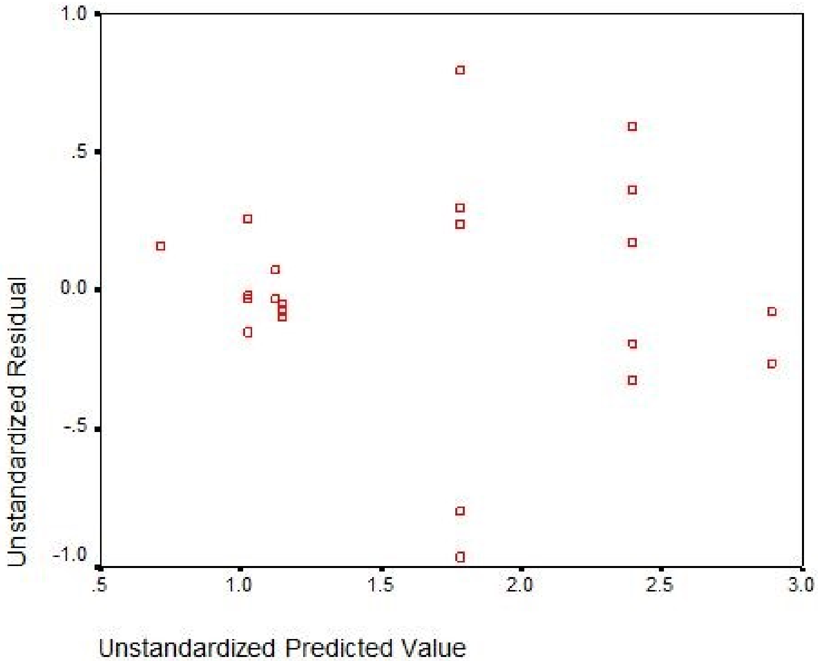
a. Predictors: (Constant), VAR00007, VAR00006, VAR00003, VAR00004, VAR00002, VAR00005

b. Dependent Variable: VAR00001

4. Residual analysis

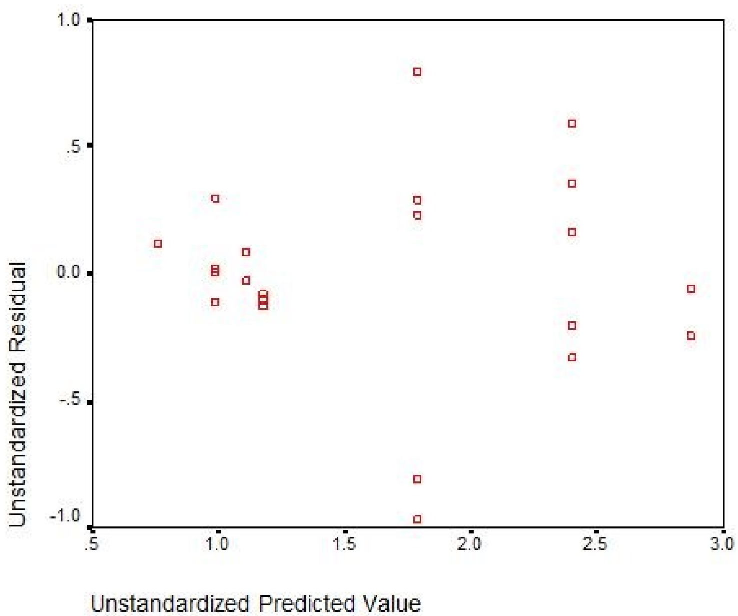
- **Testing of error variance**: One of the assumptions of ordinary least squares regression is that the variance of the residuals is homogenous across the predicted value known as **homoscedasticity**. If the developed model is well fitted then there should no specific pattern to the residual plot. If the variance of the error or residual is non–constant then the residual variance is said to be heteroscedastic. From the graph of MLR model M1-M4 we can see that the variance around zero is scattered and thus we can say that homoscedasticity is satisfied.

M1



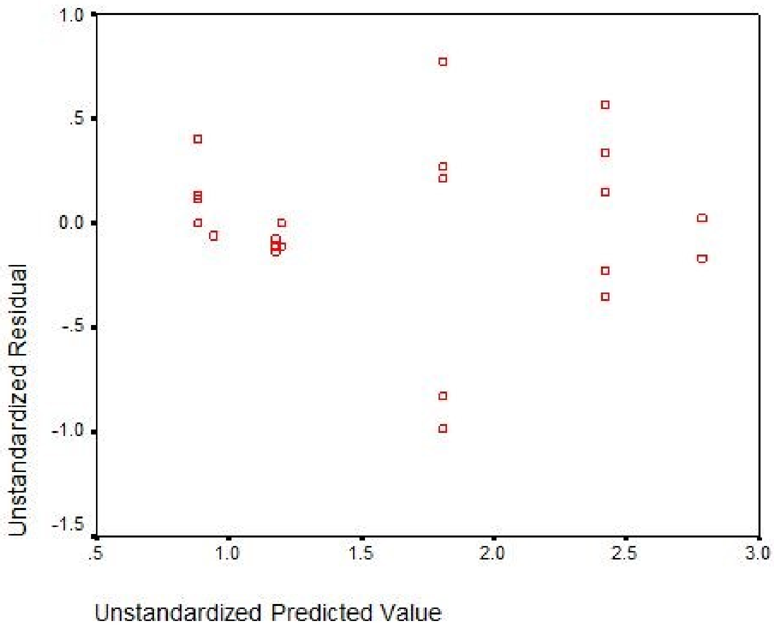
1.0
.5
0.0
-.5
-1.0
Unstandardized Residual
.5
1.0
1.5
2.0
2.5
3.0
Unstandardized Predicted Value

M2

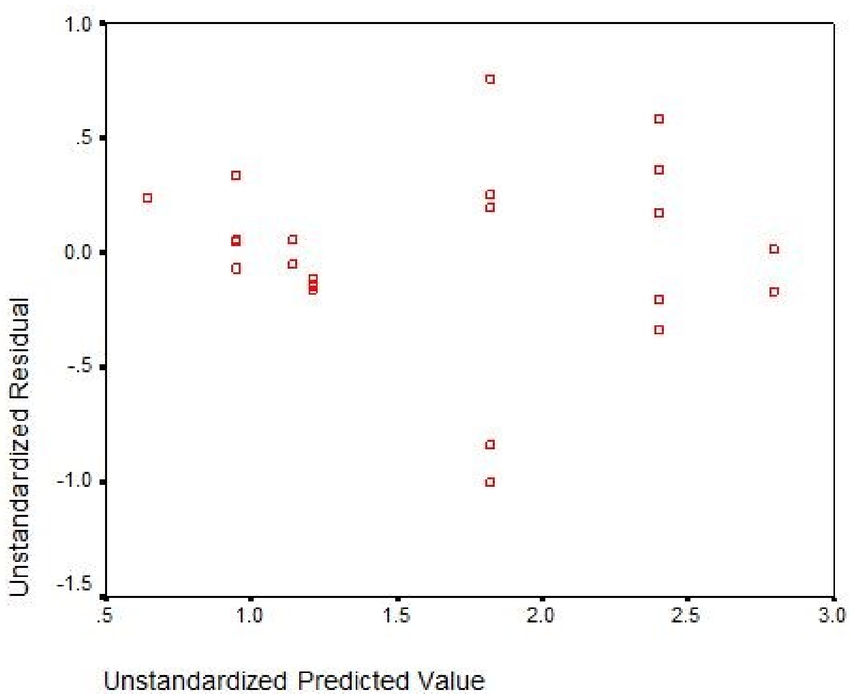


1.0
.5
0.0
-.5
-1.0
Unstandardized Residual
.5
1.0
1.5
2.0
2.5
3.0
Unstandardized Predicted Value

M3

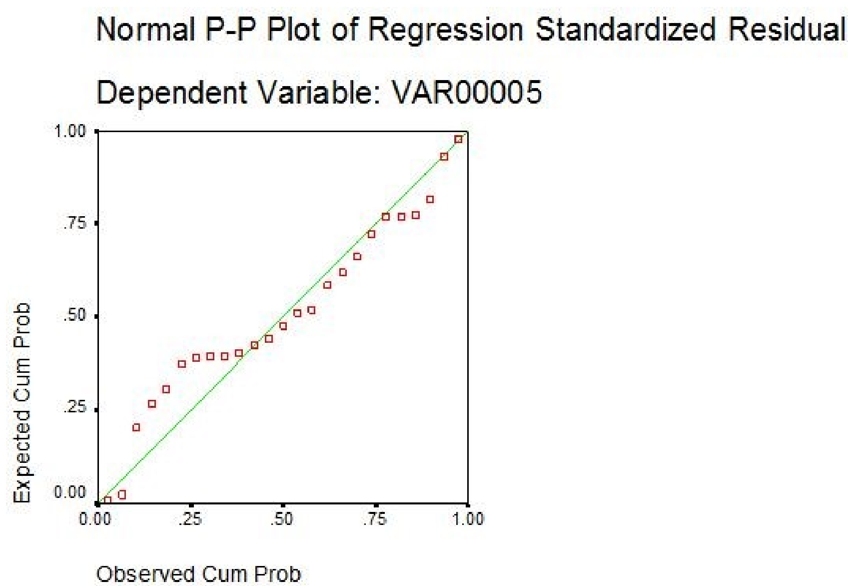

1.0
.5
0.0
-.5
-1.0
-1.5
Unstandardized Residual
.5
1.0
1.5
2.0
2.5
3.0
Unstandardized Predicted Value

M4

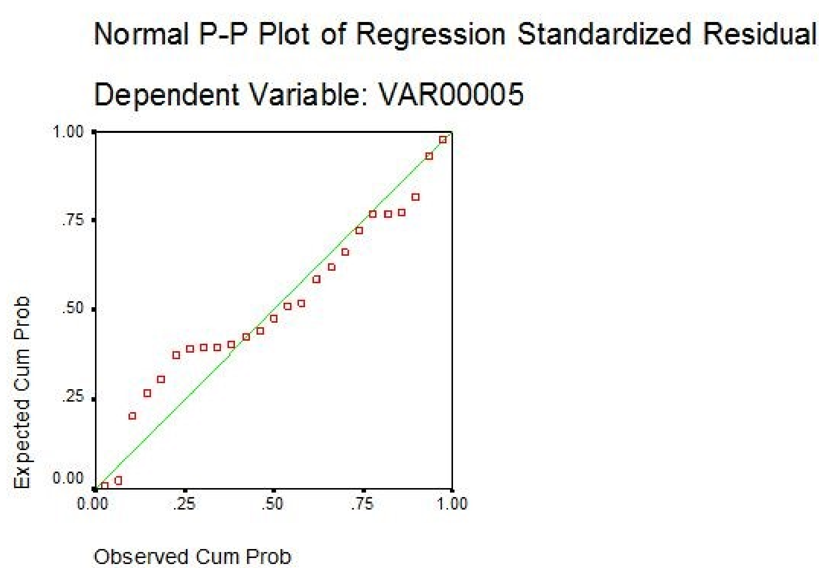

1.0
.5
0.0
-.5
-1.0
-1.5
Unstandardized Residual
.5
1.0
1.5
2.0
2.5
3.0
Unstandardized Predicted Value

➤ **<u>Test of normality of residuals:</u>** Linear regression has a misconception that the the outcome has to be normally distributed but the fact is that the residual needs to be normally distributed. Therefore we have performed probability plot. The plot compares the observed cumulative distribution function (CDF) of the standarized residual to the expected CDF.

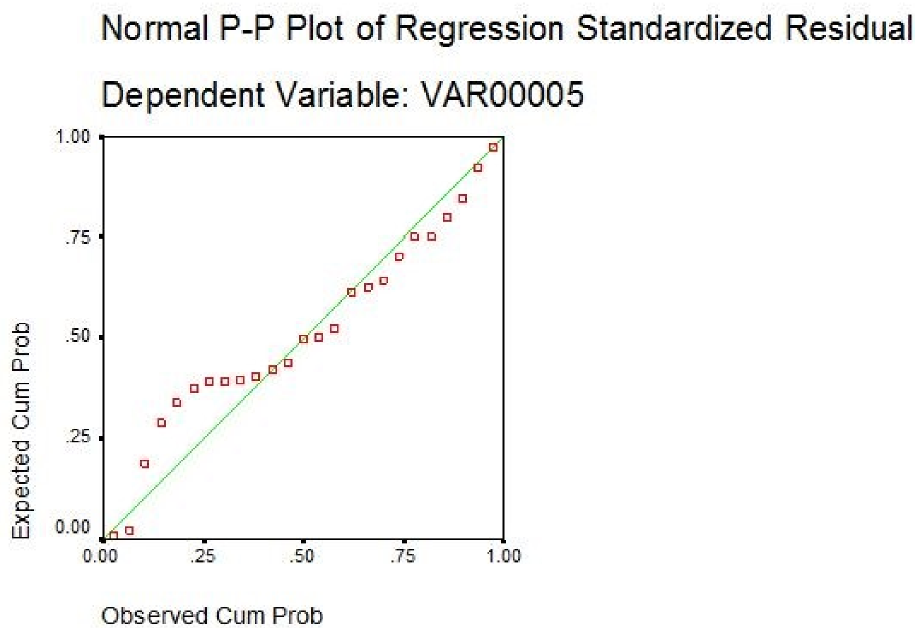
M1



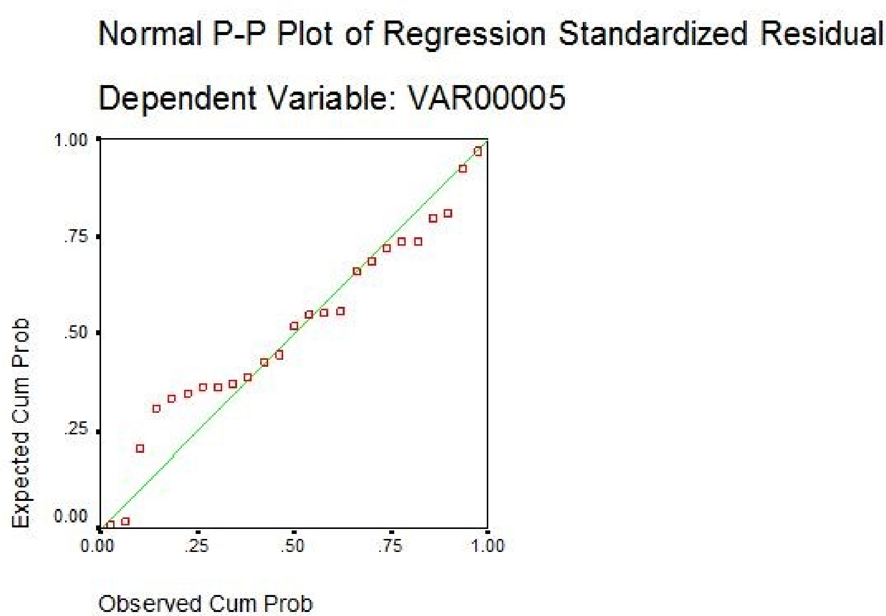
M2



M3


Normal P-P Plot of Regression Standardized Residual
Dependent Variable: VAR00005
Expected Cum Prob
Observed Cum Prob
0.00
.25
.50
.75
1.00

M4


Normal P-P Plot of Regression Standardized Residual
Dependent Variable: VAR00005
Expected Cum Prob
Observed Cum Prob
0.00
.25
.50
.75
1.00

5. **<u>Inter correlation between descriptors</u>**

Correlation is a statistical technique that shows how strongly two variables are related to each other. The correlation coefficient should be within -1 to 1 to be accepted. We have checked the correlated between the descriptors in each developed MLR models and found that the descriptors in each model were not correlated.

M1

	Metals_SumIP	SuMElectrons_Active_M_SM	solubility
Metals_SumIP	1	-0.063	-0.008
SuMElectrons_Active_M_SM	-0.063	1	-0.462
solubility	-0.063	-0.462	1

M2

	Metals_SumIP	cat_rad	solubility
Metals_SumIP	1	0.402	-0.008
cat_rad	0.402	1	-0.408
solubility	-0.008	-0.408	1

M3

	cat_rad	solubility	oxidation degree of metal (n_oxy_M)
cat_rad	1	-0.408	0.045
solubility	-0.408	1	-0.58
oxidation degree of metal (n_oxy_M)	0.045	-0.58	1

M4

	D1_Metals	cat_rad	solubility
D1_Metals	1	0.558	-0.394
cat_rad	0.558	1	-0.408
solubility	-0.394	-0.408	1