Stinchcomb et al., 2016 paper results in R

Part 1: ***Data used****:* geology_basic.csv

Ran PLSR model in R using oxide weight percentages (eg., Fe2O3_wt, Al2O3_wt...) and cl_MAP and cl_MAT

Note: a) ln(Cao) was used

b) Data with zero values are not excluded from the analysis

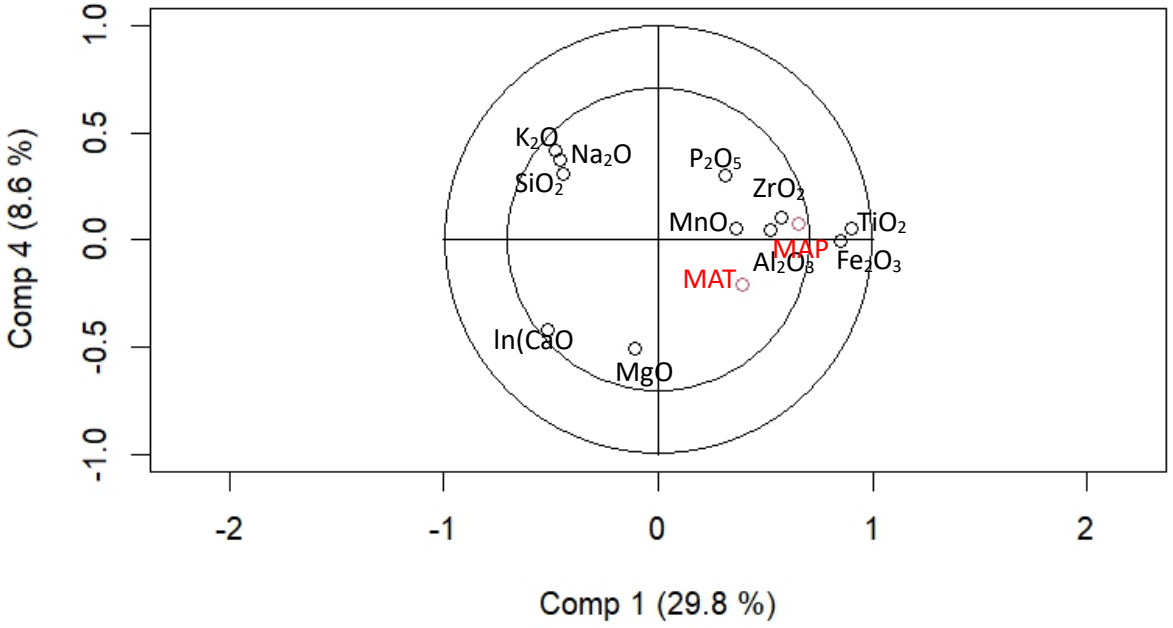
Table 1: PLSR model in R details

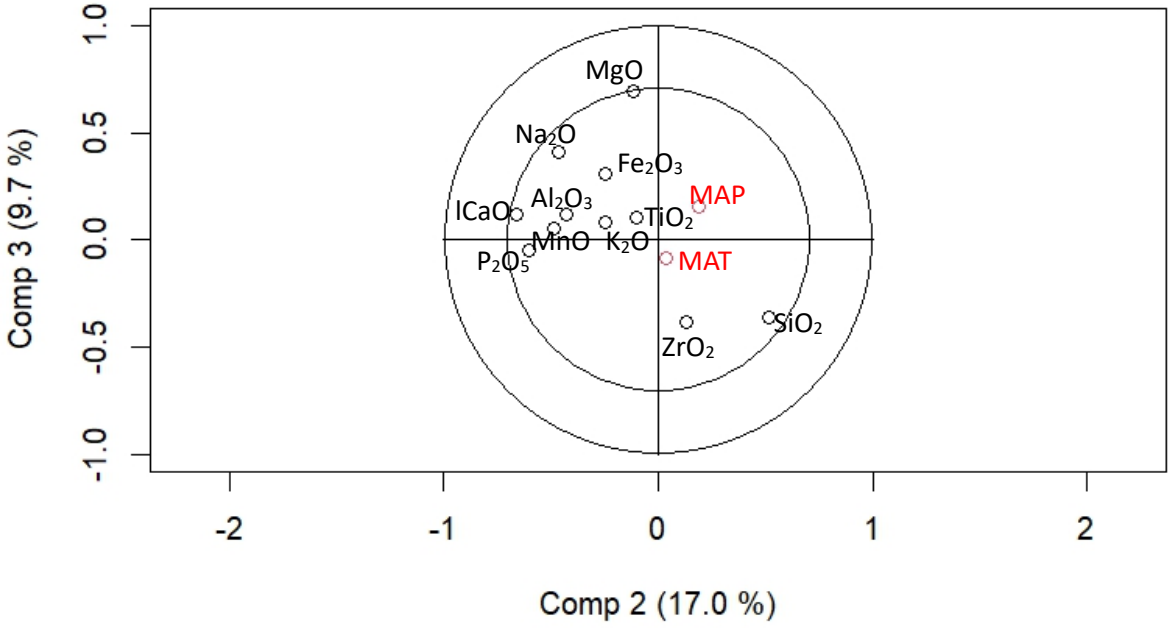
PLS algorithm	NIPALS
Response variables	2 (MAT and MAP)
Predictor variables	11 (oxides)
Missing value handling	excluded
Number of factors	4

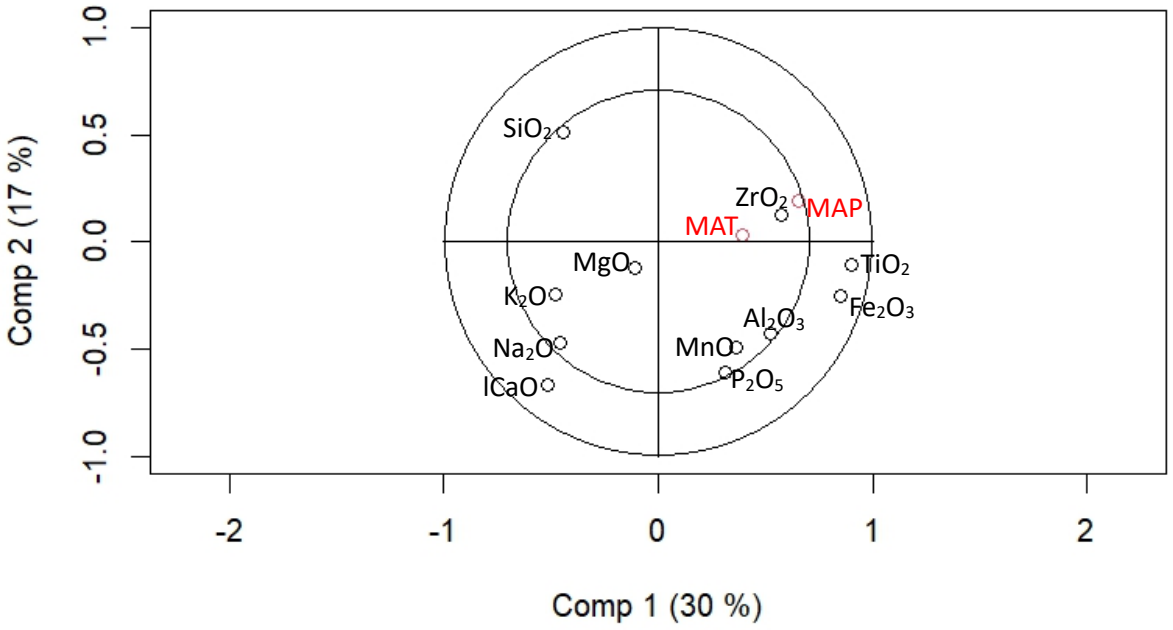
Number of observations used	685

Variance:

Percent Variance explained by the PLSR model				
Factors extracted	Predictor variables		Dependent Variables	
	Current	Total	Current	Total
1	29.77	29.77	29.35	29.35
2	17.04	46.81	1.90	31.25
3	9.72	56.53	1.62	32.88
4	8.65	65.18	2.46	35.34


Comp 4 (8.6 %)
1.0
0.5
0.0
-0.5
-1.0
-2
-1
0
1
2
Comp 1 (29.8 %)
K2O
Na2O
SiO2
P2O5
ZrO2
MnO
TiO2
MAP
Al2O3
Fe2O3
MAT
ln(CaO
MgO


Comp 3 (9.7 %)
1.0
0.5
0.0
-0.5
-1.0
-2
-1
0
1
2
Comp 2 (17.0 %)
MgO
Na2O
Fe2O3
MAP
lCaO
Al2O3
TiO2
MnO
K2O
P2O5
MAT
SiO2
ZrO2


Comp 2 (17 %)
1.0
0.5
0.0
-0.5
-1.0
SiO2
ZrO2
MAP
MAT
MgO
TiO2
K2O
Fe2O3
Al2O3
Na2O
MnO
P2O5
lCaO
-2
-1
0
1
2
Comp 1 (30 %)

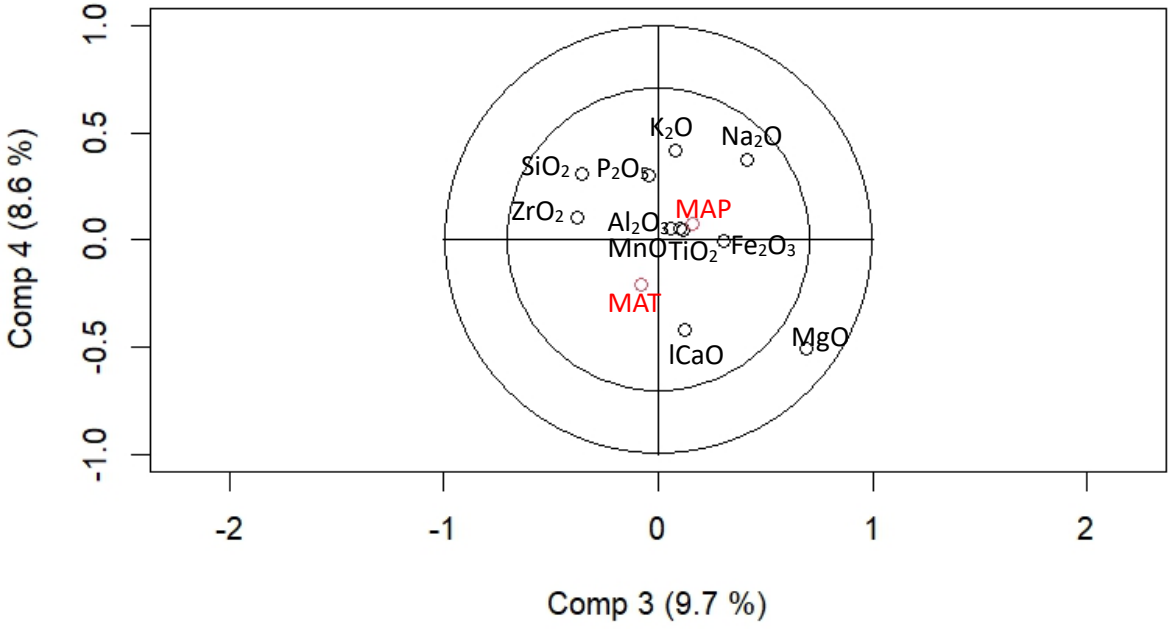

Comp 4 (8.6 %)
1.0
0.5
0.0
-0.5
-1.0
K2O
Na2O
SiO2
P2O5
ZrO2
MAP
Al2O3
MnO
TiO2
Fe2O3
MAT
MgO
lCaO
-2
-1
0
1
2
Comp 3 (9.7 %)

Table 2: Correlation coefficients, Pearson's r for components 1-4 (C1-C4), oxides (wt%), MAP and MAT loadings

	C1	C2	C3	C4	MAP	MAT
Fe2O3	0.85	-0.25	0.31	-0.01	0.55	0.30
MnO	0.36	-0.49	0.06	0.05	0.16	0.08
P2O5	0.31	-0.60	-0.04	0.30	0.11	0.05
SiO2	-0.44	0.51	-0.36	0.31	-0.25	-0.25
TiO2	0.90	-0.10	0.10	0.05	0.60	0.33
ZrO2	0.57	0.13	-0.38	0.10	0.34	0.12
Al2O3	0.52	-0.43	0.12	0.05	0.27	0.19
Na2O	-0.49	-0.46	0.41	0.37	-0.30	-0.36
K2O	-0.48	-0.24	0.08	0.42	-0.33	-0.33
MgO	-0.11	-0.12	0.69	-0.51	-0.02	-0.06
ln(CaO)	-0.51	-0.66	0.12	-0.42	-0.47	-0.15
MAP	0.66	0.19	0.16	0.08		
MAT	0.39	0.03	-0.08	-0.21		

Note: Running PLSR using molar concentrations or weight percentages didn't show any variation in the results, however, there is a little variation observed in correlation coefficients running the model using CaO vs ln(CaO)

- The correlation loading plot between components (or factors) 1 and 4 was closer to the Stinchcomb et al., 2016 paper
- Component 1 is similar to the regressor 1 in Stinchcomb paper
- The Pearson correlation coefficient between MAP, MAT, and component 1 is similar to the Stinchcomb paper but the rest are a bit different

Results from Stinchcomb et al., 2016

TABLE 3

Correlation coefficients, Pearson's r, for Regressors1-4 (R_1-R_4), geochemical oxide, and MAP and MAT loadings.

Variable	R_1	R_2	R_3	R_4	MAP	MAT
Fe_2O_3	0.85	0.22	-0.19	0.28	0.55	0.33
MnO	0.36	0.36	-0.36	0.02	0.16	0.08
P_2O_5	0.31	0.39	-0.46	-0.05	0.11	0.05
SiO_2	-0.47	0.11	0.66	-0.38	-0.25	-0.25
TiO_2	0.89	0.19	-0.03	0.07	0.60	0.33
ZrO_2	0.54	0.34	0.36	-0.39	0.34	0.12
Al_2O_3	0.52	0.32	-0.34	-0.01	0.27	0.19
Na_2O	-0.48	0.62	-0.24	0.33	-0.30	-0.36
MgO	-0.10	0.00	-0.19	0.74	-0.02	-0.06
K_2O	-0.51	0.60	0.01	-0.08	-0.33	-0.33
ln(CaO)	-0.49	0.03	-0.74	0.12	-0.47	-0.15
R_1	1.00	0.00	0.00	0.00	0.65	0.41
R_2	0.00	1.00	0.00	0.00	0.00	-0.25
R_3	0.00	0.00	1.00	0.00	0.21	-0.08
R_4	0.00	0.00	0.00	1.00	0.16	-0.05
MAP	0.65	0.00	0.21	0.16	1.00	0.33
MAT	0.41	-0.25	-0.08	-0.05	0.33	1.00

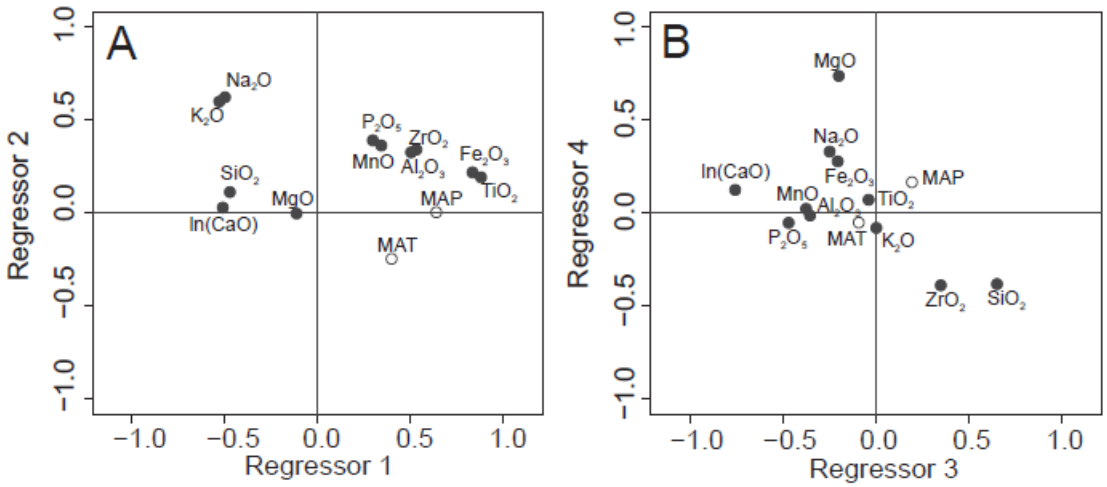


Fig. 5. Correlation loading plots for independent (geochemical oxides) and dependent (MAP and MAT) variables by regressor. Regressors are plotted against one another to emphasize differences in element loadings and correlation with MAP and MAT. The loadings show the amount of variability in each geochemical oxide that is accounted for by the Regressors, 1–4, and the degree to which the Regressors explain the response variables, MAP and MAT. See table 2 for corresponding *r* values. Note that *r* does not directly reflect oxide or climate loadings, **P'** or **C'**, that were determined in an 11-dimensional space using Partial Least Squares Regression (PLSR).

Part 2: ***Data used***: ppm1_data.sas7bdat from the sas data uploaded by Jack in Github

Ran PLSR model in R using oxide molar concentrations (eg., Fe2O3_mol, Al2O3_mol...) and cl_MAP and cl_MAT

Note: a) Cao_mol was used in this section instead of ln(CaO_mol)

b) Data with zero values are not excluded from the analysis

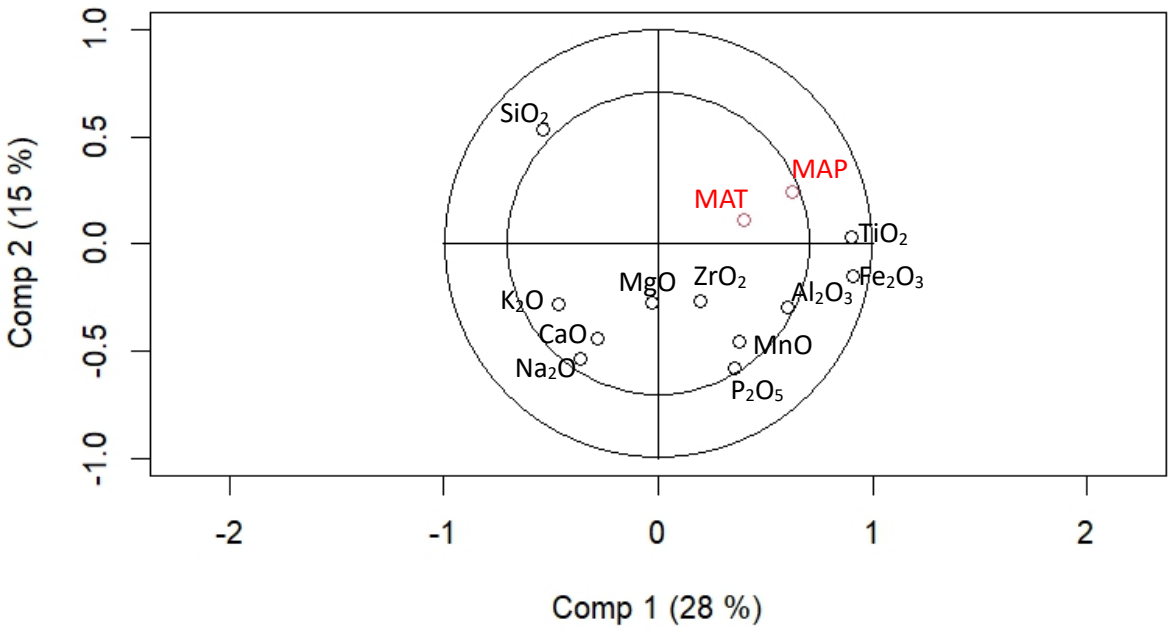
Table 1: PLSR model in R details

PLS algorithm	NIPALS
Response variables	2 (MAT and MAP)
Predictor variables	11 (oxides)
Missing value handling	excluded
Number of factors	4

Number of observations used	685

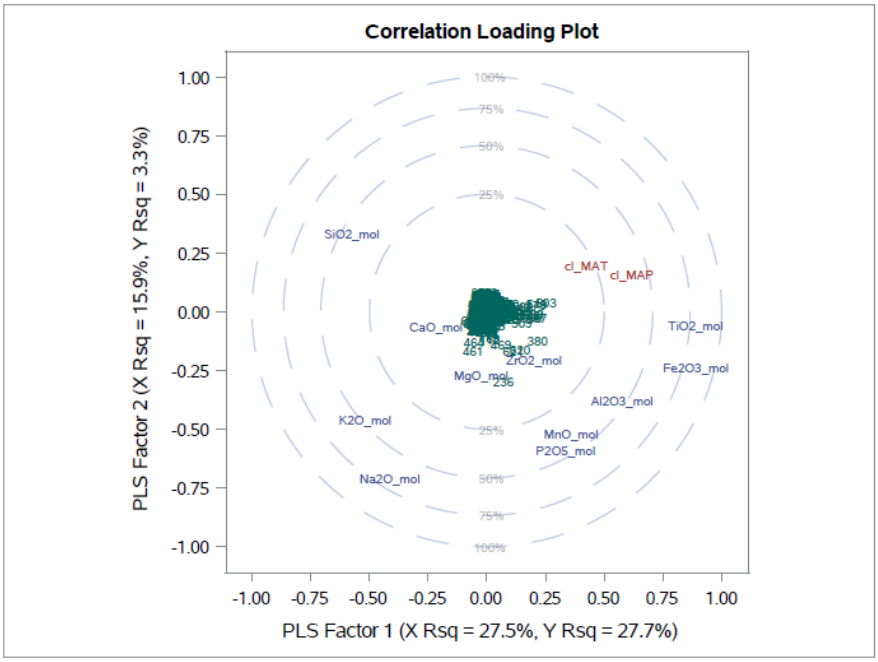
Variance:

Percent Variance explained by the PLSR model				
Factors extracted	Predictor variables		Dependent Variables	
	Current	Total	Current	Total
1	27.62	27.62	27.16	27.16
2	14.89	42.51	3.54	30.70
3	8.79	51.30	1.23	31.94
4	6.48	57.78	1.43	33.37

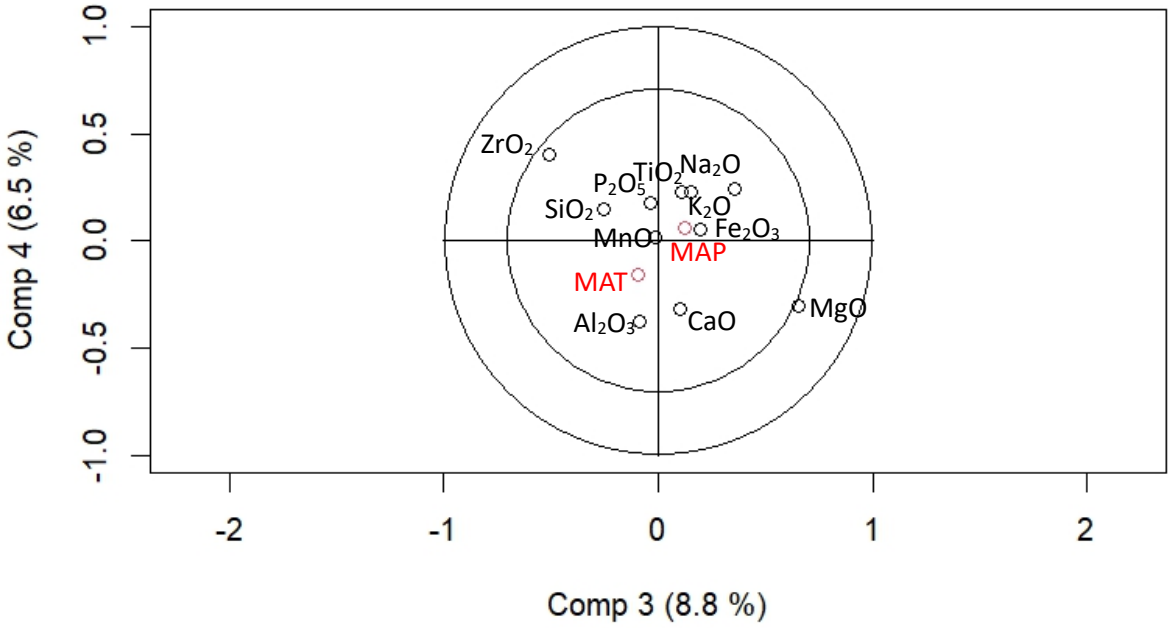


SoilGeoChem
PLS

The PLS Procedure



- The correlation loading plot ran in R using sas data was similar to the sas generated plot, expect for CaO. The variance (Table 1) is also similar to sas one
- I used the same data that was used in sas to generate the above figure in R



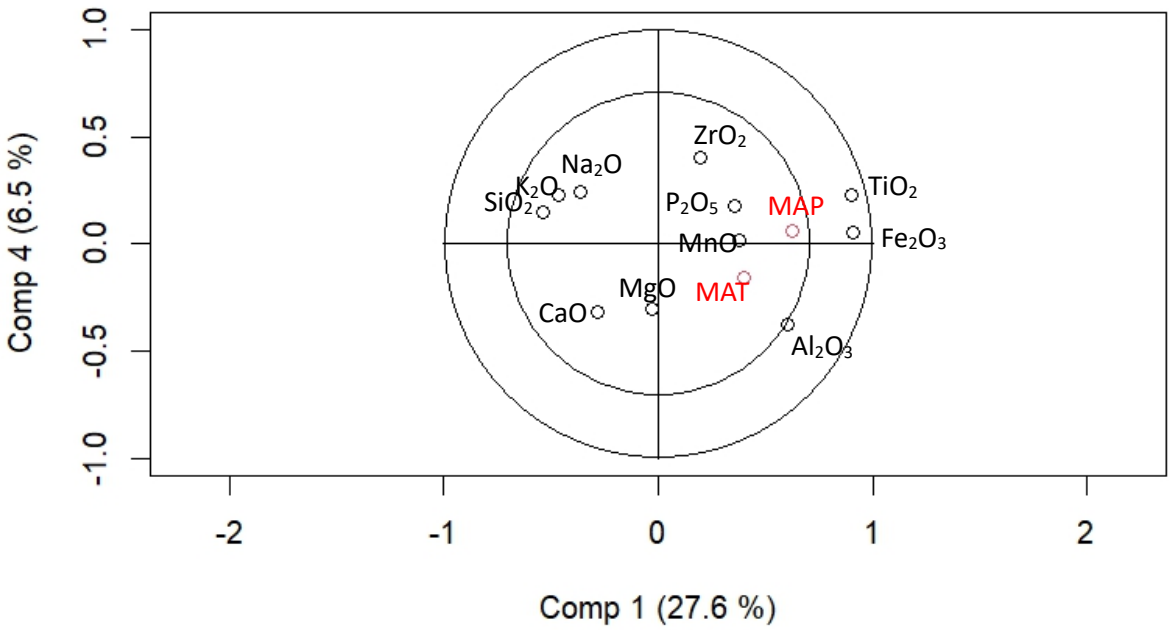


Table 2: Correlation coefficients, Pearson's r for components 1-4 (C1-C4), oxides, MAP and MAT loadings

	C1	C2	C3	C4	MAP	MAT
Fe2O3	0.91	-0.17	0.26	0.06	0.55	0.30
MnO	0.38	-0.48	-0.05	0.004	0.16	0.08
P2O5	0.35	-0.59	-0.11	0.19	0.11	0.05
SiO2	-0.54	-0.50	-0.35	0.36	-0.25	-0.25
TiO2	0.90	0.009	0.07	0.11	0.60	0.33
ZrO2	0.19	0.21	-0.52	0.06	0.31	0.10
Al2O3	0.60	-0.39	-0.01	0.06	0.27	0.19
Na2O	-0.36	-0.55	0.26	0.34	-0.30	-0.36
K2O	-0.47	-0.35	-0.07	0.51	-0.33	-0.33
MgO	-0.02	-0.19	0.69	-0.37	-0.02	-0.06
CaO	-0.28	-0.38	0.07	-0.68	-0.28	0.02
MAP	0.62	0.24	0.13	0.06		
MAT	0.40	0.11	-0.09	-0.16		