Using JMP® software to develop glass formulations for high-level nuclear waste conditioning: data visualization, statistical analysis and predictive models

Perret D.[(1)], Bardez-Giboire I.[(1)], Dussossoy J.L.[(1)], Bousquet N.[(1)], Baudet F.[(2)]
[(1)] CEA Marcoule, DTCD/SECM/LDMC-LCV, BP 17171, F-30207 Bagnols-sur-Cèze, France
[(2)] JMP France, Domaine de Grégy, Grégy-sur-Yerres, F-77257 Brie Comte Robert Cedex, France

1. Introduction

Developing glass formulations for high-level nuclear waste conditioning has been one of the most challenging issues for more than 40 years in the nuclear industry. Long-term storage of radioactive waste requires the stabilization of the waste into a form which will neither react nor degrade for extended periods of time. One way to do this is through vitrification which is a particularly attractive immobilization way. Indeed, some borosilicate glasses have the capability to incorporate a significant amount of nuclear waste while keeping a high chemical durability.
Glass formulations for the vitrification of high-level nuclear waste have been under investigation at the French Atomic Energy Commission (CEA) for many years. Beside the complexity of its formulation, nuclear glass also needs to meet requirements which are specific to the industrial vitrification process. For example, viscosity, density, electrical and thermal conductivities, and of course, long term durability of the glass, are properties that have to be perfectly understood and controlled.
As a consequence, we continuously have to deal with huge amounts of data, including formulation data (glass compositions), physical and chemical property data, and data related to the vitrification process. JMP software has been recently selected to be implemented in our R&D teams who develop nuclear glass formulations. As it is described in this paper, JMP provides very useful and easy-to-use tools, which enable, among others, the comparison of glass composition domains having high degree of complexity, or the development of property-to-composition predictive models.

2. Glass formulation

As traditional glassy materials, nuclear glasses exhibit an amorphous structure, where the radioactive elements coming from the waste are part of the glass network, and linked to the glass-forming elements with true chemical bonds (Figure 1 and Figure 2). The high degree of complexity of nuclear glasses is especially related to the number of chemical elements that have to be incorporated into the glass structure. Typically, the final glass formulation has to contain all the fission products coming from the reactions occurring in the nuclear reactor (Figure 4). At the end, more than thirty elements are present into the glass (Figure 5). As a result, the comparison of glass compositions is very complex and requires specific statistical and geometrical tools. The JMP graphical platform provides very useful and easy-to-use tools, such as the Scatterplot Matrix, Ternary Plot or Mixture Profiler platforms, which enable the visualization and the analysis of large amounts of formulation data. For example, the Scatterplot Matrix platform is very convenient to help us analyzing the final quality of the glass, by showing which chemical elements play a key role in the crystallization mechanisms (Figure 6). For this purpose, the Partition tool in the JMP Modeling platform can also be used very efficiently.

But according to us, one of the main interests of using JMP relies on its powerful statistical analysis platform, which enables the comparison of glass composition domains having high degree of complexity. For this purpose, PCA (Principal Component Analysis), Cluster and Dendrogram platforms are very relevant. When you need to compare composition domains for materials that contain three components only, it is possible to use classical Ternary Plots (Figure 7 and Figure 8). But in our case where glasses have p components with p higher than 10, and where all these components have individual and relational constraints, the domains you would need to compare have the shape of convex polyhedra in a p-dimensional space. Therefore the data visualization becomes more complicated.

PCA is a mathematical method used to reduce the number of variables, by creating a smaller set of independent (uncorrelated) variables, which will account for most of the variance in the observed variables. For our purpose where we need to compare glass composition domains, we use the PCA platform of JMP as a data projection tool in a 2-dimensional space. Then it becomes possible to visualize whether a composition domain is close, or far, from another domain, or whether two domains are secant, for instance (Figure 9).

The hierarchical clustering platform is also an efficient tool enabling a relevant analysis of a high number of glass compositions. Generally speaking, cluster analysis includes a broad suite of techniques designed to find groups of similar items within a data set. Partitioning methods divide the data set into a number of groups predesignated by the user. Hierarchical cluster methods produce a hierarchy of clusters from small clusters of very similar items to large clusters that include more dissimilar items. Hierarchical methods usually produce a graphical output known as a dendrogram or tree that shows this hierarchical clustering structure. The dendrogram lists all of the samples and indicates at what level of similarity any two clusters were joined. The x-axis is some measure of the similarity or distance at which clusters join. For our application, the distance between the glass compositions is calculated with traditional mathematical methods, like the Euclidian distance for example.

JMP has the capability to propose several methods for hierarchical clustering, like the Average Linkage (Figure 10), who tends to produce clusters with the same variance, the Centroid Method (based on the Euclidian distance calculation), who has the advantage of being more robust to outliers than most other hierarchical methods, or the Ward's method, who tends to produce clusters with roughly the same number of observations. Therefore by using the JMP clustering platforms, we are able to make a relevant use of the thousands of nuclear glasses we have been elaborating for many years.

3. Glass properties and database

The methodology we use for optimizing high-level waste (HLW) glass formulation involves collecting and generating a property-composition database, and relating these properties to glass composition. This methodology has to be very robust for increasing the efficiency and decreasing the cost of the final vitrification industrial process. Therefore our studies are focused on building property-to-composition predictive models. The general method is to use the experimental data points we have generated from a set of glass formulations inside a wide composition domain, in order to predict the behaviour of the glass at any point of the composition domain. This can be efficiently carried out by using Fit Model and Stepwise platforms. For this application, the JMP Mixture Response Surface platform is convenient and Stepwise method is easy-to-use. The Predictive Model report (Figure 11) gives every useful piece of information: actual by predicted plot, fit summary, parameter estimates with standard errors, residual plots, prediction profiler and PRESS (Predicted Error Sum of Square) value. PRESS values are very important data because they are the criteria we use to quantify the quality of the prediction given by our models. Unfortunately, JMP 9 cannot use the PRESS

value as a criterion for Stepwise regression. It would be for us a significant improvement for the next versions of JMP.
The predictive model report also enables to save many statistical data into columns in the tables, such as the Hats, which are important values for identifying the observations which have a large effect on the outcome of the fitting model (Figure 12). Therefore, even if we do not use JMP to build our Mixture Design of Experiments, we use it for the whole statistical analysis of the DOEs. For this purpose, the JMP capability of importing external data tables makes a real difference against using most of the traditional DOE softwares with which you need to build up the full DOE before starting the statistical analysis.
The PCA platform as described before can be used with glass composition variables and with property variables as well. Then it becomes possible to make a 2-dimensional projection of all the composition and property variables for identifying the main correlations (Figure 13). This is of major importance in order to understand the physical and chemical mechanisms occurring into the glass structure.
Finally, a property-composition database is being generated by our glass formulation teams. The database will contain data related to the vitrification process conditions at the laboratory scale, glass compositions and all the physical and chemical properties. Because JMP has the capability to connect to external database, we will use it for the statistical analysis of data gathered in previous research programs over the past decades, in order to build more robust predictive models.

4. Conclusion

JMP software has been recently selected to be implemented in our R&D teams who develop nuclear glass formulations. JMP provides relevant tools, which enable, among others, the comparison of glass composition domains having high degree of complexity, or the development of property-to-composition predictive models. At a laboratory scale, JMP helps us to understand the physical and chemical mechanisms occurring into the glass structure. At an industrial scale, the methodology we use for optimizing HLW glass formulation involves collecting and generating a property-composition database, and relating these properties to glass composition. JMP is a precious tool to make this methodology very robust, increasing the efficiency and decreasing at the end the cost of the final vitrification industrial process.

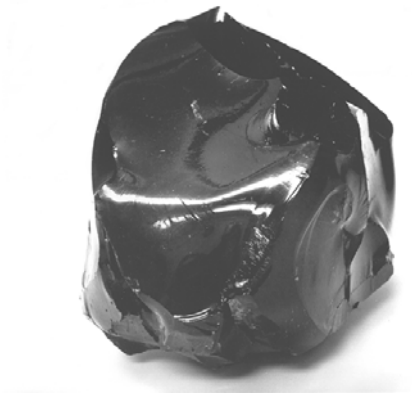


Figure 1: Nuclear glass sample (inactive sample)

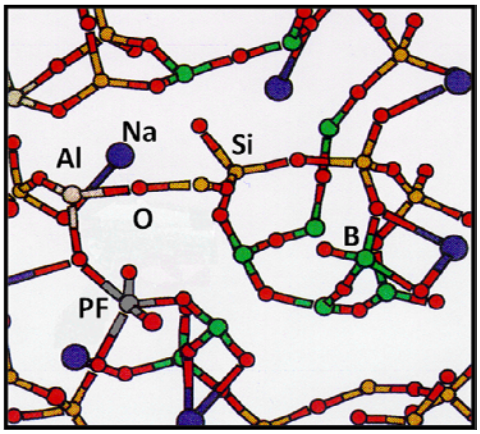


Figure 2: Amorphous structure of nuclear glass network. Fission product elements (PF) are part of the network.

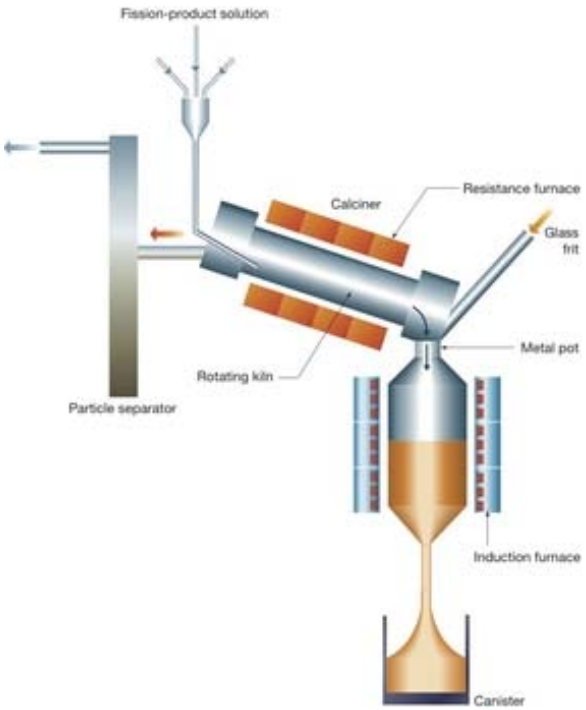


Figure 3: Vitrification process for High-Level Waste (HLW) conditioning
Reference: S. Naline et al., "Vitrification 2010-A Challenging French Vitrification Project to Retrofit a Cold Crucible Inductive Melter at the La Hague Plant" – 10382 – presented at the Waste Management conference in Phoenix, Arizona, 7-11 March 2010

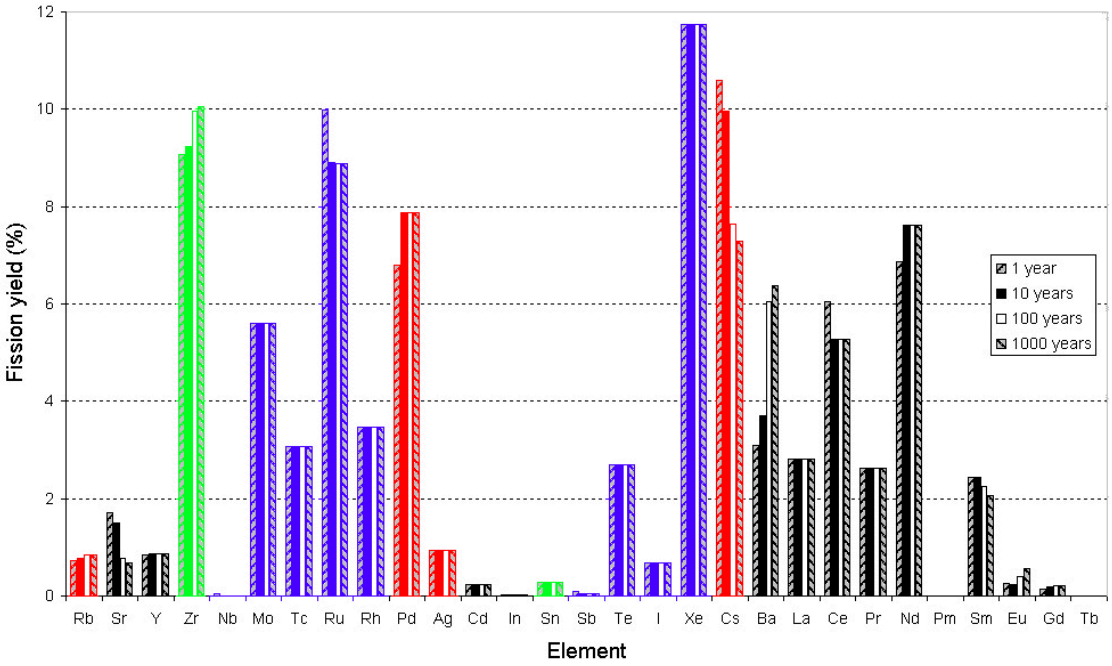


Figure 4: Fission products obtained from thermal neutron fission of U235
Reference: http://en.wikipedia.org/wiki/Fission_product_yield

Description of the three composition domains: variation ranges (wt%) and relational constraints

	R7T7	AVM	VRZ
SiO_2	42.40–51.68	38.5–46	40.8–55.6
B_2O_3	12.40–16.50	16–19.5	1.0
Na_2O	8.10–11.00	5–18.8	
Al_2O_3	3.60–6.60	9–12.5	9–13.7
CaO	4.0	0.2	19.7–24.6
Nd_2O_3	0.5–2.6	0.1–0.9	1–6.5
ZrO_2	2.0–4.9	0.1–1.0	2–12
TiO_2			6.2–15.5
MgO		2.5–7.5	
$Fe_2O_3 + NiO + Cr_2O_3$	0.4–5.6	2.8	
F		0–1.8	
P_2O_5	0.5	0–1.7	
$FP + Act + MoO_3 + Gd_2O_3 + Ag_2O$	3.53–17.95	0–10.2	
Fines	0.01–6.85		
ZnO	2.5		
Li_2O	2.0	0.4	
SO_3		0.1	
Cl		0.1	
CdO		0.5	
Relational constraints	$3.01 < SiO_2/B_2O_3 < 3.47$ $7.0 < FP + Act + Fines < 18.0$ $FP + Act > Fines$ $SiO_2 + B_2O_3 + Al_2O_3 > 60.0$ $Fe_2O_3/NiO = 7.09$ $Fe_2O_3/Cr_2O_3 = 5.73$	Frit constant $MgO + Al_2O_3 < 18.5$ $Al_2O_3 < 3 * [(Na_2O + Li_2O) - 0.278SiO_2]$ $FP + Act + MoO_3 + Gd_2O_3 + Ag_2O > 0.5 MgO$	$TiO_2–ZrO_2 < 8$ $ZrO_2–TiO_2 < 1$

Oxide breakdown (wt%) compared with reference glass

wt%	R7T7		AVM
	FP + Act	Fines	FP + Act + ···
SrO	0.34		0.21
ZrO_2	1.70	0.47	0.96
MnO_2	0.30	0.08	0.31
Cs_2O	1.10		0.71
BaO	0.61		0.34
Y_2O_3	0.20		0.12
La_2O_3	0.92		0.56
Ce_2O_3	0.95		0.60
Nd_2O_3	1.63		0.93
Pr_2O_3	0.45		0.28
SnO_2	0.02		0.02
Sb_2O_3	0.00		0.01
TeO_2	0.23		0.14
ThO_2	0.31		0.11
UO_2	0.05	0.01	0.70
MoO_3	1.36	0.39	0.75
Ag_2O	0.03		0.10
RuO_2	0.63	0.36	
Rh	0.12	0.05	
Pd	0.33	0.10	
CdO	0.03		
Gd_2O_3			0.60

'Fines' include platinum-group metals and metallic particles, FP refers to fission products.

Figure 5: Examples of nuclear glass composition domains
Reference: P. Frugier et al., Journal of Nuclear Materials, 346 (2005) 194-207

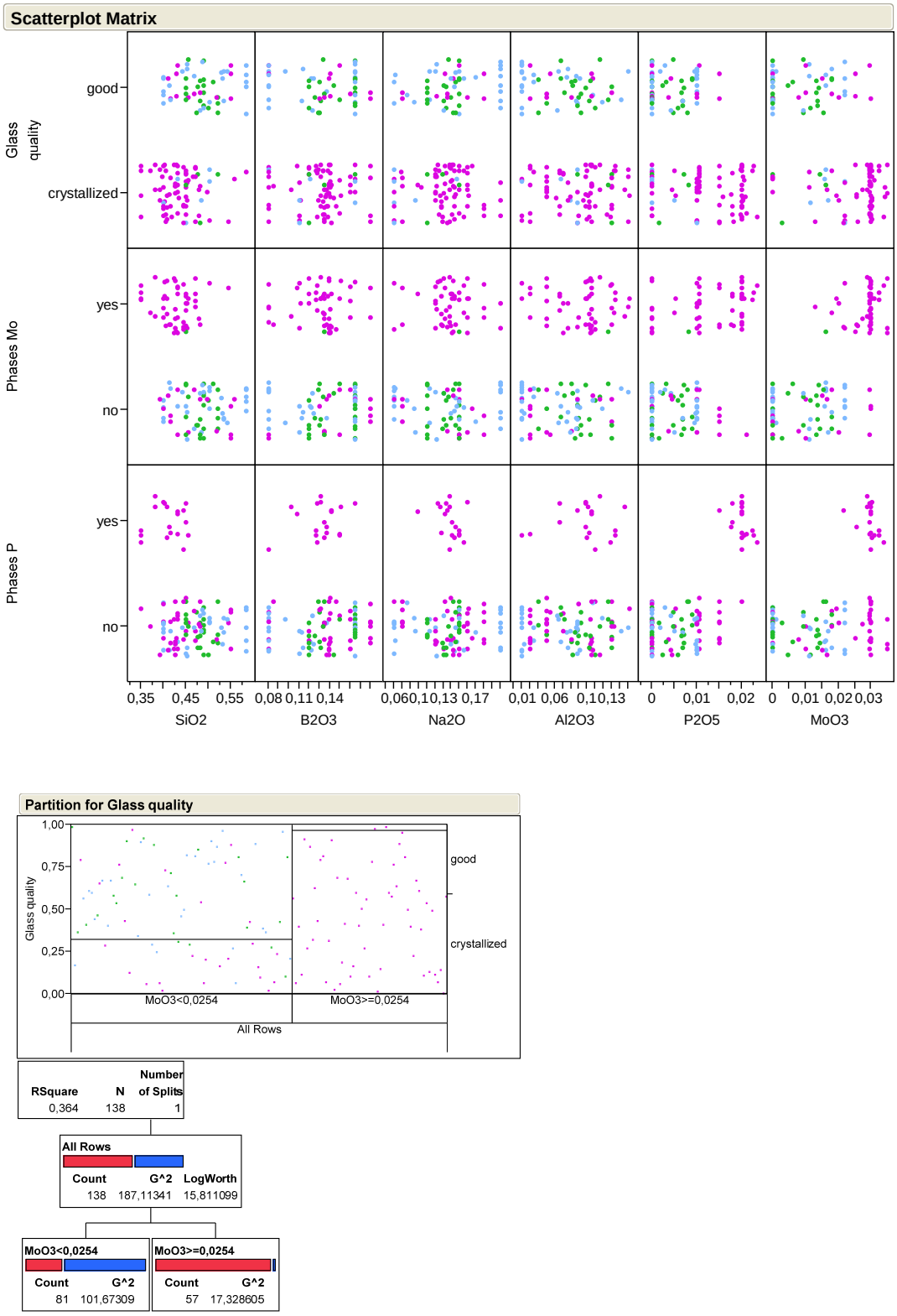


Figure 6: The Scatterplot Matrix platform enables a relevant visualization of glass formulation data. The influence of the composition on the final quality of the glass can be analyzed in order the get a better understanding of the physical and chemical mechanisms occurring into the glass structure. The Partition platform can be used as well.

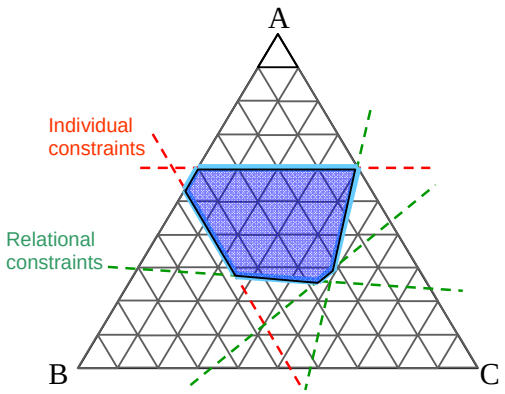


Figure 7: Individual and relational constraints in a three-component formulation, using ternary plot.
Constraints give the formulation domain boundaries. Three-component formulation domains have the shape of two-dimensional convex polyhedra.

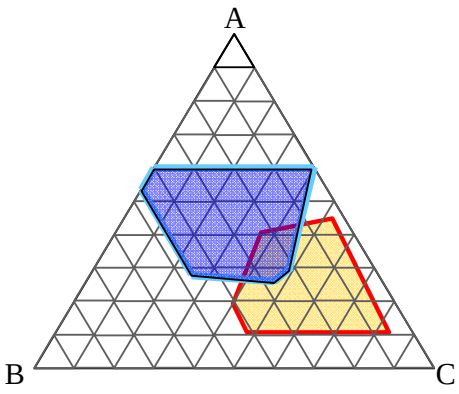


Figure 8: Comparison of composition domains in a two-dimensional space.
Ternary plots cannot be used anymore in the case of p-component formulation domains, with p higher than 3.

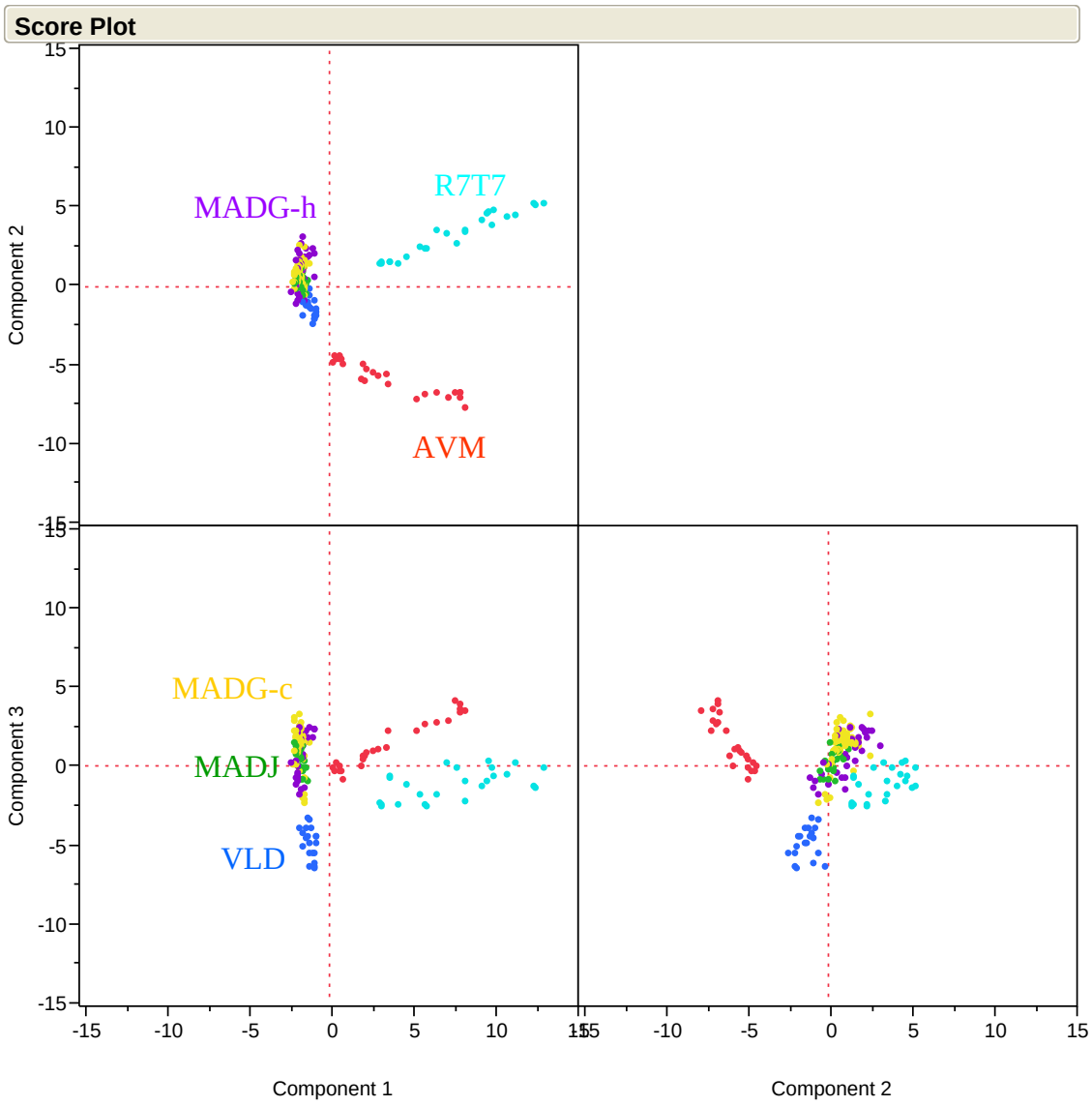


Figure 9: Principal Component Analysis (PCA) using glass formulation data coming from 6 composition domains.
The analysis is done with composition data as variables, no property data are included. The score plot shows a projection of the composition data using the 3 first principal components.

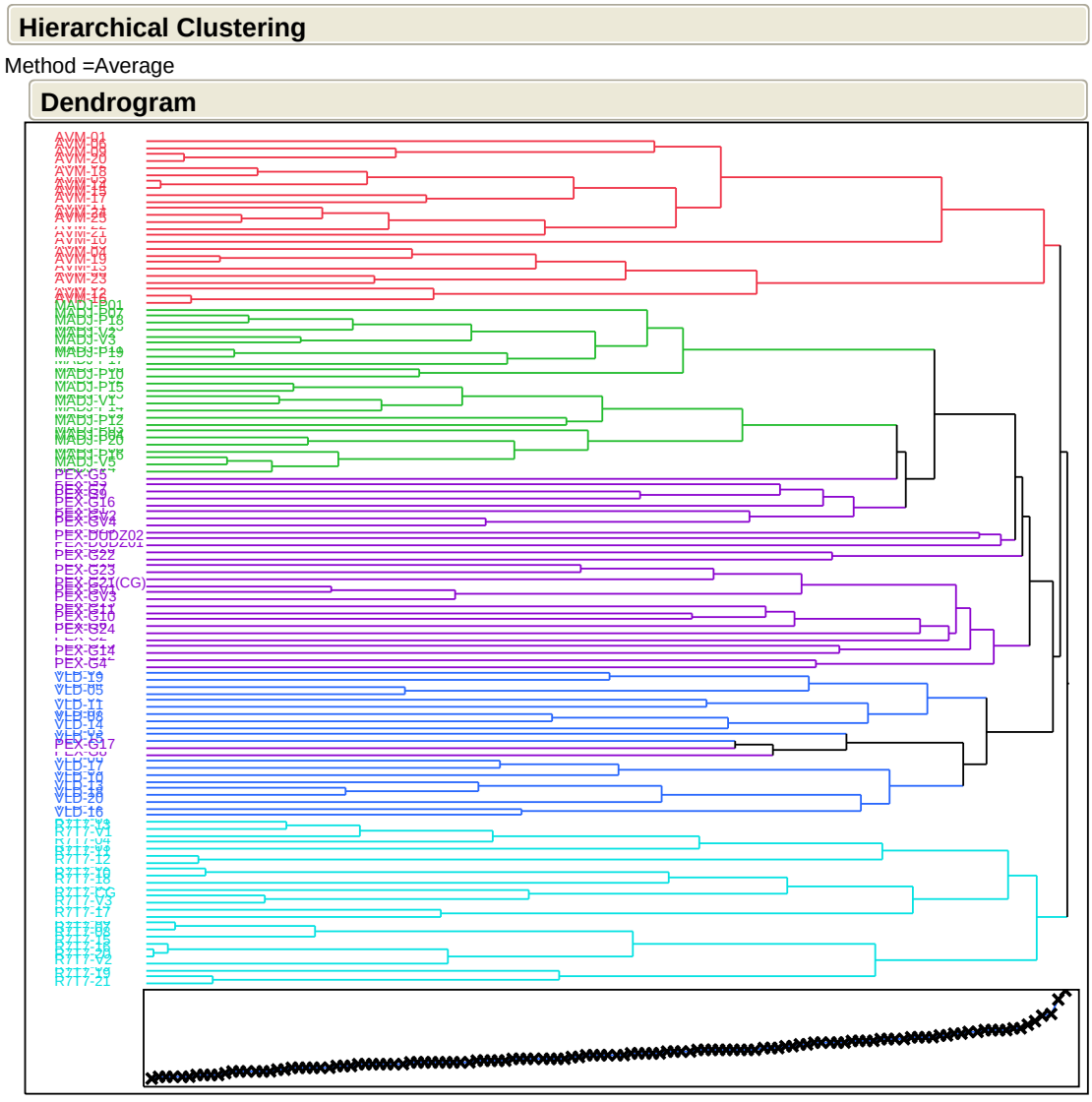
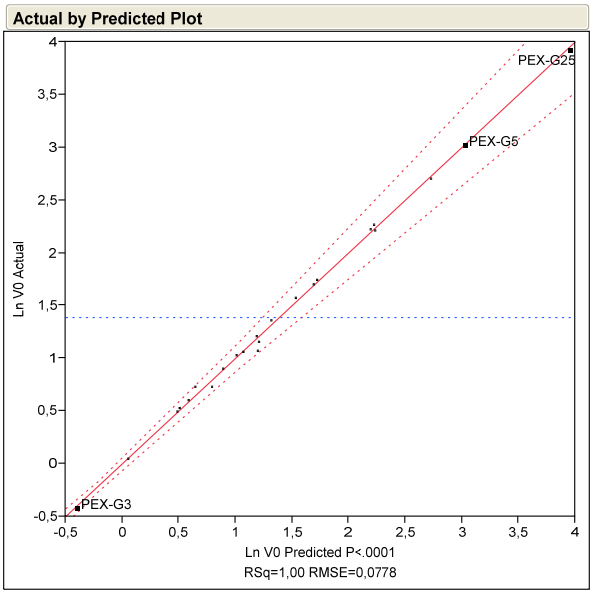
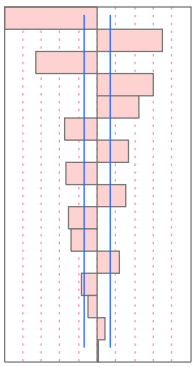


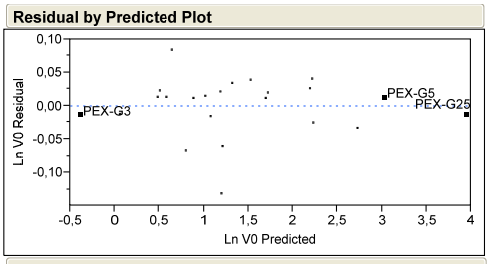
Figure 10: Hierarchical Clustering gives a dendrogram showing distance between glass compositions.
5 domains of glass compositions are analyzed. The distance on the horizontal x-axis gives an idea about the Euclidian distance between the compositions.

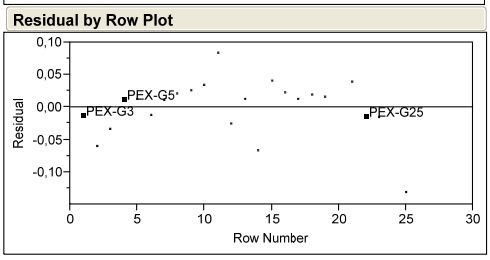


Sorted Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Si	-4,553498	0,272791	-16,69	<,0001*
Li	40,024865	3,445746	11,62	<,0001*
Ca*Ce	-1537,857	139,4036	-11,03	<,0001*
Ce	60,103954	5,947338	10,11	<,0001*
Ca	37,562478	5,002885	7,51	0,0001*
Mo	-28,73137	4,964843	-5,79	0,0007*
B*Na	136,71395	24,53483	5,57	0,0008*
ETR	-19,89196	3,572994	-5,57	0,0008*
Zr*Mo	649,60078	124,5004	5,22	0,0012*
TRA	-16,90788	3,364108	-5,03	0,0015*
Al	-7,348866	1,594729	-4,61	0,0025*
Fe	6,7020361	1,652943	4,05	0,0048*
Zr	-8,541538	3,025097	-2,82	0,0256*
Na	-3,644185	2,162095	-1,69	0,1358
P	15,62458	11,50135	1,36	0,2165
B	1,1432305	3,72074	0,31	0,7676







Summary of Fit

RSquare	0,998085
RSquare Adj	0,993982
Root Mean Square Error	0,077784
Mean of Response	1,385612
Observations (or Sum Wgts)	23

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	15	22,074491	1,47163	243,2302
Error	7	0,042353	0,00605	Prob > F
C. Total	22	22,116844		<,0001*

Tested against reduced model: Y=mean

Press

Press	Press RMSE
0,5026635534	0,15115668

Prediction Profiler

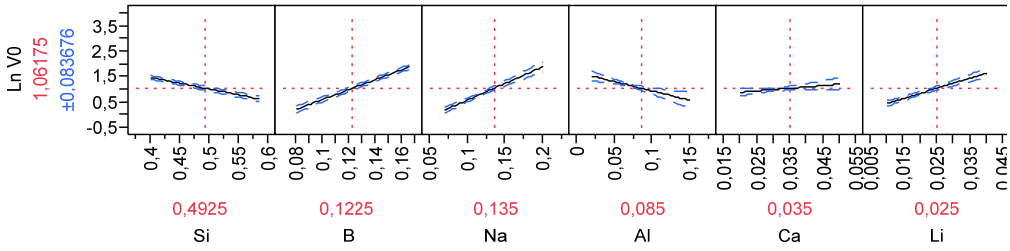


Figure 11: Statistical report from the Fit Model platform.

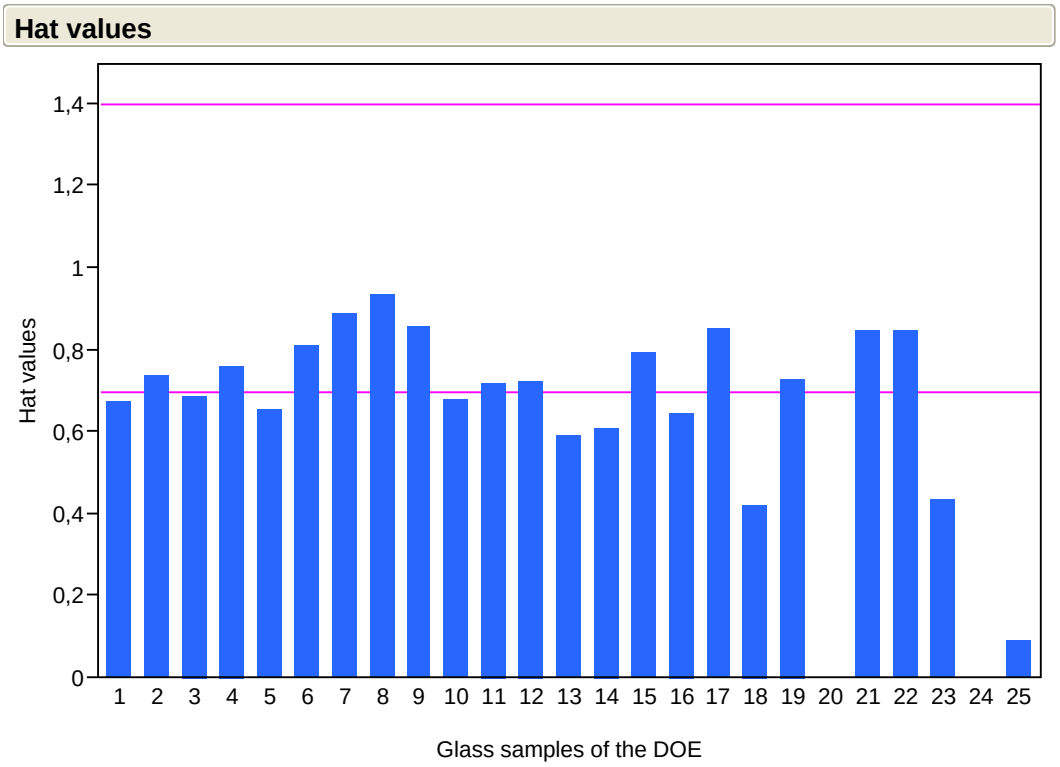


Figure 12: Hat values h_{ii} are important data for identifying the observations which have a large effect on the outcome of the fitting model.
average(h_{ii}) = p/n (p=number of model parameters ; n = number of experiments) = 0.7 for this model
A run with h_{ii} greater than 2 times the average is generally regarded as having high leverage, i.e., compared to the other runs, it is an outlier in the independent variable space.
Composition of glass sample #25 is at the center of the domain, explaining why the leverage of this point is much lower than other samples of the DOE.

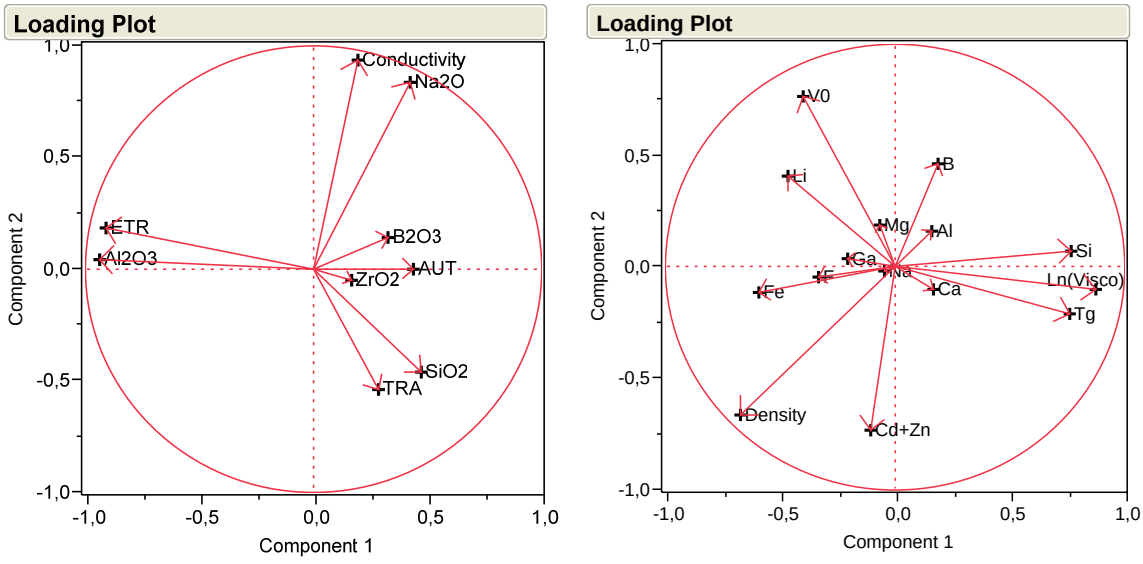


Figure 13: Principal Component Analysis (PCA) using both composition and property data as variables.