Title: Deep learning prediction of material properties

**Authors :** **Charles Le Losq1,****2,3,\*, Andrew Valentine2 , Bjorn O. Mysen3, Daniel R. Neuville1**

**Affiliations :**

1Institut de physique du globe de Paris, Université de Paris, Paris 75005, France

2 Research School of Earth Sciences, Australian National University, Canberra 2601, Australia

3 Geophysical Laboratory, Carnegie Institution for Science, Washington D.C. 20001, U.S.A.

\*Correspondence to: [lelosq@ipgp.fr](mailto:lelosq@ipgp.fr)

*We want to dedicate this manuscript to the memory of G.N. Greaves, which sadly passed away this year. His contributions to the field of material and glass sciences will be remembered.*

**Abstract:** (121/125 w)

Knowledge of the tryptic relationship between chemical composition, molecular structure and macroscopic properties of materials like minerals or glasses is fundamental in many disciplines, including metallurgy, Earth and material sciences. Several avenues allow prediction of this relationship, but usually this task remains very challenging. Deep learning represents a new tool to perform such task, allowing one to link experimental results, models, and even different theories. Focusing on alkali aluminosilicate melts, of prime importance for volcanology and glass sciences, we show how a deep neural network leverages experimental data and theoretical knowledge for prediction of melt and glass properties, including viscosity, fragility, density, or even glass Raman signals. Such approach provides a new, exciting way to build models for various applications.

**One sentence summary** (96/125 characters): Deep learning bridges the gap between data and models for practical material property prediction

**Main Text:**

The intrinsic properties of materials are governed by their chemical composition, and the associated atomic/ionic structure. Understanding this connection is fundamental to material engineering and manufacturing, and also to addressing many key questions across physics, chemistry and the geosciences. Information may be drawn from several sources, particularly experimental observations and thermodynamic or molecular dynamics (MD) simulations, or a combination of these (e.g. *1*, *2*). Techniques such as CALPHAD calculations (*3*) and MD simulation (*4*) have proved successful for some simple materials (such as binary alloys), or under specific conditions (e.g. very high temperatures in MD). However, they are not yet able to effectively predict the properties of more complex materials, such as oxide glasses or minerals within the pressure-temperature regimes relevant to industry or the Earth sciences. In these cases, thermodynamic models often offer an intermediate way of predicting macroscopic properties, but the experimentalist still has to link thermodynamic variables to data, and this often turns out to be challenging. We propose that machine learning can serve as the “middle man”, allowing scientists to make sense of their data and to put them in the context of different models and theories.

As a concrete example, we consider the links between the chemical composition, structure and physical properties of aluminosilicate melts and glasses (e.g., see the review of *5*). These materials are central to the Earth and material sciences: they form the liquid part of magmas, with melt viscosity directly determining whether volcanic eruptions have effusive or explosive character (*6*), and, as glasses, they are so important to modern technology and life that we could be said to be living in The Glass Age (*7*). A number of different theories exist that attempt to connect the structure of these melts to their properties. These often involve variables that are either difficult or impossible to measure directly, such as the glass configurational entropy *Sconf*. In this publication, we employ a deep neural network to represent such variables within the physical theory, and use this to infer their relationship to composition. For simplicity, we focus on compositions in the K2O-Na2O-Al2O3-SiO2 system, as a fairly complete (albeit sparse) dataset is available (Fig. S1).

There is a vast landscape of theories (and their associated set of equations) describing the viscous flow of liquids, supported by different arguments depending on use case (e.g. *11*). Within each theory, viscosity is modelled as dependent on a certain subset of the melt’s physical properties. For example, the Adam-Gibbs model (8) represents the viscosity () as a function of temperature (*T*) and composition (*x*)

, (1)

where *Ae* represents the behaviour in the high-temperature limit, is a term proportional to the to the energy barriers opposed to the cooperative molecular re-arrangements of melt structure upon viscous flow, *Sconf* and *Cpconf* are the melt configurational entropy and heat capacity respectively, and *Tg* is the melt glass transition temperature. Alternatively, one might adopt the MYEGA model, which gives the relationship as

(2)

where and are defined as before, and where represents the glass fragility. Many other models may also be found in the literature, and a few are given in the Supplementary Material.

In order to make use of any of these models, it is necessary to determine values for the composition-dependent physical parameters such as and . However, in many cases these are difficult to measure directly; instead, they must be inferred by fitting the viscosity model to experimental observations. This brings its own disadvantages: the value for inferred using the Adam-Gibbs model may differ from that for obtained with the MYEGA model. Which is ‘correct’? Each of these models is built upon its own set of assumptions, and for some applications it may be evident that one choice is more appropriate than others. Often, however, there is no clear rationale for adopting one model instead of another; one simply requires good predictive performance (i.e., the ability to accurately forecast the viscosity of a melt given its composition and temperature). In such cases, it may be appropriate to employ trans-theoretical modeling, combining predictions from multiple theories in pursuit of better overall behaviour.

In this paper, we use a deep neural network to simultaneously predict the values of a variety of different physical properties of aluminosilicate melts, given information about composition (Fig. 1). The chosen physical properties are those required to predict viscosity within five theoretical frameworks: the Adam-Gibbs, MYEGA, Avramov-Milchev, Tamman-Vogel-Fulcher and free volume theories (see Supplementary Material). In addition, we predict three composition-dependent quantities that are readily observed experimentally: the quenched melt density, its raman spectrum, and the parameters necessary to predict its refractive index at a given wavelength, according to the Sellmeier equation. We train the neural network using a suite of measurements of viscosity, density, raman and optical spectra, seeking best average performance across all theories. In particular, it should be noted that where quantities (such as ) appear in multiple theoretical frameworks, our approach provides one value that is optimal across all models.

After training (see details in Supplementary Text), the model allows systematic predictions of melt viscosity using the Adam-Gibbs theory, as shown in Figure 2A where model predictions and measurements are compared for melts with ~66 mol% SiO2 and varying Al/(Na+K) and K(K+Na). Other theories can also be used, like the Free Volume (FV) theory that yields a precision similar to the AG theory (Fig. 2B,C). Semi-empirical and empirical equations, like the famous Tamman-Vogel-Fulcher equation, yield a good precision too (Supplementary Text and Figure S3). In all cases, predictive errors are lower than 0.50 log Pa·s, as shown by root-mean-square-errors (RMSE) ~ 0.45 log Pa·s on the *testing* data subset (Table S3). Errors are slightly higher in the supercooled domain (viscosity > 107 Pa·s) for all theories (Fig. S3) because, in this domain, melt entropy (eq. 1) varies in a complex manner depending on melt composition (see discussion) and this results in large, non-linear viscosity variations difficult to reproduce. Interestingly, based on only numerical performance, the different viscosity models yield similar precisions (Table S3 and Figure S3) and it may seem difficult to discriminate a better one (Fig. 2B,C; see also Table S3). However, the AG theory relates measurable thermodynamic variables to melt viscosity (*9*) and performs as well as, or better than others equations. Consequently, present results confirm that this theory seems well appropriate for aluminosilicate melts. Aside this and more generally, the obtained results clearly indicate that deep learning allows practical trans-theoretical modeling, a feature extremely interesting in many areas.

***Thermodynamic and static property predictions***

Beyond trans-theoretical viscosity predictions, the present model offers the possibility to predict aluminosilicate melt thermodynamic properties: known viscous *Tg and* *Sconf(Tg)* are predicted within 16 K and 0.8 J mol-1 K-1, respectively (RMSE on testing subsets; Fig. S4). Glass properties are also successfully predicted, as shown by RMSE on testing density and refractive index data subsets of 0.009 g cm-3 and 0.005, respectively. Confident in the model predictive ability, we can now use it to explore how parameters like *Sconf(Tg)*, a driver of melt rheology (eq. 1), vary with melt composition in the entire glass-forming domain of the Na2O-K2O-Al2O3-SiO2 diagram.

Variations in *Sconf(Tg)* with melt composition are non-monotonic, with maxima close to, or on the silicate joint (alkali-silica, Al-free), while in general a decrease in *Sconf(Tg)* is observed upon addition of aluminum (Fig. 3A,B). Increasing the fraction of MAlO2 (M=Na, K) decreases *Sconf(Tg)*, the effect being more linear and pronounced for potassic compositions (Fig. 3B). Mixing alkalis result in different effects as a function of the M/Al ratio of the melt (Fig. 3C,D). For silicate compositions, the Mixed Alkali Effect (MAE) (*15*) is observed: a maximum in *Sconf(Tg)* is generally observed as Na and K mix, except at very low (< 0.55) or very high (> 0.95) silicate contents. At high Al concentrations, the variations in entropy with the Na/K ratio are closer to a linear trend, indicating a different MAE (*2*).

Similarly to melt entropy, the melt glass transition temperature *Tg* vary largely as a function of melt composition (Fig. 4A,B). This allows large changes in melt dynamic properties with only limited variations in fragility (Fig. S6). We finally note that variations in glass density and refractive index with composition are also smooth and monotonic, both parameters depending mostly on the glass SiO2 fraction, and in lesser extents on its M/Al as well as Na/K ratios (Fig. 4C,D,E,F). Those parameters are actually correlated between themselves, as visible in corner plots (see supplementary Jupyter Notebooks).

***Raman spectra: a glimpse into melt structure***

Raman spectra were used as a way to perform multitask learning (Supplementary Text) and to introduce structural information in the artificial neural network. The dataset we have is limited (48 compositions), such that we did not aim at having excellent predictions, but we simply wanted the network to capture the general Raman signal variations as a function of glass composition. This goal is achieved: predicted and measured *RRaman*, the ratio between signals assigned to intertetrahedral vibrations (below 670 cm-1) and intratetrahedral vibrations (between ~860 and 1300 cm-1) compare well (Fig. 5A), and the median of the relative mean absolute deviation between predicted and measured Raman spectra (Fig. S7) is of 22 % on the validation *DRaman* subset. The model thus captured the general link between glass composition and glass structure as sampled by Raman spectroscopy, but remains affected by some evident predictive errors (Fig. S7), particularly at high silica concentrations where the dataset is particularly scares (Fig. S1). Better results could be obtained by improving the coverage of the Raman spectral database in the future.

**Discussion**

**The developed framework offers unique opportunities to objectively explore the links between chemical, structural, thermodynamic and dynamic parameters in materials.** Such task is traditionally done by humans, and this situation is far from being ideal as we are influenced by various biases (*16*). For example, because of the known link between melt structure and properties, one can propose that melt viscosity could be predicted solely from *RRaman* (*17*), with the benefit that this value is measured via a cheap, rapid to perform, non-damaging, spatially-resolved spectroscopy that can be easily deployed *in situ*. However, this is forgetting the importance of ionic mixing effects on the thermodynamic properties of melts (*15*, *9*, *13*). Indeed, Raman spectra are only a spectroscopic picture of the AlO4-SiO4 polyhedral network topology, which is barely affected by ionic mixing effects (*18*). Therefore, it should not be possible to predict melt viscosity from only Raman spectra. Our results confirm this: in the investigated system, while there is a general non-linear correlation between *RRaman* and *Tg* for instance (Fig. 5B; Spearman correlation coefficient *rs* = 0.98), *Sconf(Tg)* and melt fragility *m* display complex variations with *RRaman* (Fig. 5C,D). Such complexity further explains why no general model of *Sconf(Tg)* or *m* has been proposed to date.

**The trans-theoretical character of the model allows adopting a new vision about a problem like melt viscous flow and glass transition.** For aluminosilicate melts, it allows observing the link between two theories: Adam-Gibbs (AG) and Free Volume (FV). In the FV theory, solid-like and liquid-like molecular cells are distinguished and separated by a critical volume *v\**, and viscous flow occurs via cooperative molecular movements between liquid-like cells. In the AG theory, viscous flow occurs via cooperative motions of molecular segments of a size *z\*(T),* characterized by an intrisic entropy *Sc\**. The two theories thus share some common philosophical background: viscous flow is assumed to occur via some sort of cooperative movements of molecular entities in the melt. This background can be retrieved when diving into the details of the parameters of eq. 1 and 2. Indeed, *BCG* embeds some structural information because it depends on *v\**:

*BCG = v\* zo, (5)*

where *zo* is an adjustable parameter. Similarly, the ratio *Be/Sconf(Tg)* embeds molecular subunits lengthscale information as

*Be/Sconf(Tg) = [ Δµ z\*(Tg) ] / R, (6)*

with *Δµ* the energy barriers opposed to the rearrangement of molecular subunits of size *z\*(Tg)*, and R the perfect gas constant. Now let’s consider *v\** and *z\** as structural parameters embedding information about the volume or lengthscale of cooperative molecular regions. In this case, they should directly depend on melt or glass structure. This is confirmed by the fact that both *BCG* and *Be/Sconf(Tg)* correlate very well with *RRaman* (Fig. 5E,F, *rs* = 0.987 and 0.985 respectively), which provides information about the network connectivity; the higher *RRaman* is, the higher the interconnection between polyhedral units, thus the higher the 3D network topology *(19*)*.* In detail, this implies that it actually should be possible to develop a free-volume version of the AG theory, as it has been proposed (*20*, *21*). More generally, the links between *BCG,* *Be/Sconf(Tg)* and *RRaman* support the hypothesis that melt viscous flow occurs when a critical molecular lengthscale is reached. This lengthscale can be determined from Raman spectra (Fig. 5E,F) and strongly influences the glass transition temperature *Tg* (Fig. 5B). In detail, entropic effects (like the excess of entropy resulting from the MAE) also affect *Tg* (*18* and references cited therein) but their influence remains limited compared to that of the polyhedral network topology (Fig. 5B). On the other hand, such entropic effects strongly control the rate at which supercooled melt viscosity changes as a function of T, or in other terms, the melt fragility *m* because (Fig. S5)*.* Actually, we may consider that such mixing phenomena are allowed by the non-ergotic nature of melts, their inhomogeneities being frozen into glasses below *Tg*. Following from this idea, it is possible to relate *m* to ongoing density fluctuations inherited from dynamic heterogeneities formed at supercooled T, i.e. Tliquidus < T < Tg (*22*). The nature of such heterogeneities is expected to change largely with melt composition, as we can recognize different cases for network organization as described by the Random Network for simple AX2 glasses like SiO2 (*23*), the Modified Random Network for silicate liquids (*24*) and the Continuous Compensated Random Network (*25*, *2*) for aluminum-rich aluminosilicates, each one presenting distinct and different extent of heterogeneities at medium and long range order with various associated cationic mixing within such heterogeneities that drive large and complex changes in glass configurational entropy (Supplementary Text) and hence viscosity (Fig. 2A).

**The presented approach is flexible, improvable, and transposable.** The present model, which successfully links chemistry, structure and properties of aluminosilicate melts and glasses, is a proof a concept, and is meant to change and evolve in the future. Different improvements will consist in, for instance, (i) extending it to compositions containing cations like Ca, Mg, Ti or Fe, (ii) including other equations to predict melt viscosity, or (iii) adding other spectroscopic information like 29Si and 27Al nuclear magnetic resonance spectra. At the moment, the principal limitation of the model resides in the limited sets of experimental data available for training. The combination of different methods (early stopping, dropout, multitask learning, bagging) allowed successful training of the model on small datasets, but a significant number of compositions are still needed to obtain good results (more than 80 in the present case, see Supplementary Text and Fig. S2). This may represent a problem for including compositions for which available data are very sparse, like aluminate melts. In addition, entropy variations due to mixing of different alkali and alkaline-earth cations (e.g., Ca-K, Mg-K…) remain not well known, and this may be another difficulty to solve in the future. Despite such potential problems, and considering existing data on simple melts as well as geologic compositions, extension of the present model to compositions pertinent for geologic and industrial melts is realistic. More importantly, the present approach can be readily transposed to other problems in the fields of chemistry, physics, geo- and material sciences where only part of the problem can be solved via know equations, and other parts linking many different observations on an object like a material or a system are fully unknown but can be learn from data using deep learning. The present work proves that this is possible on small datasets, and can bring invaluable information about long-standing problems.

References and Notes. (40 citations max main text)

1. Y. Wang, T. Sakamaki, L. B. Skinner, Z. Jing, T. Yu, Y. Kono, C. Park, G. Shen, M. L. Rivers, S. R. Sutton, Atomistic insight into viscosity and density of silicate melts under pressure. *Nat. Commun.* **5** (2014), doi:10.1038/ncomms4241.

2. C. Le Losq, D. R. Neuville, W. Chen, P. Florian, D. Massiot, Z. Zhou, G. N. Greaves, Percolation channels: a universal idea to describe the atomic structure and dynamics of glasses and melts. *Sci. Rep.* **7**, 16490 (2017).

3. N. Saunders, A. Miodownik, *CALPHAD (calculation of phase diagrams): a comprehensive guide* (Elsevier., 1998).

4. D. C. Rapaport, *The art of molecular dynamics simulation* (Cambridge University Press, Cambridge, U.K., ed. 2nd, 2004).

5. C. Le Losq, M. R. Cicconi, G. N. Greaves, D. R. Neuville, in *Handbook of Glass* (Springer, 2019; https://www.springer.com/us/book/9783319937267).

6. D. B. Dingwell, Volcanic Dilemma - Flow or Blow? *Science*. **273**, 1054–1055 (1996).

7. D. L. Morse, J. W. Evenson, Welcome to the Glass Age. *Int. J. Appl. Glass Sci.* **7**, 409–412 (2016).

8. G. Adam, J. H. Gibbs, On the temperature dependence of cooperative relaxation properties in glass-forming liquids. *J. Chem. Phys.* **43**, 139–146 (1965).

9. P. Richet, Viscosity and configurational entropy of silicate melts. *Geochim. Cosmochim. Acta*. **48**, 471–483 (1984).

10. C. A. Angell, Relaxation in liquids, polymers and plastic crystals—strong/fragile patterns and problems. *J. Non-Cryst. Solids*. **131**, 13–31 (1991).

11. Y. Bottinga, P. Richet, A. Sipp, Viscosity regimes of homogeneous silicate melts. *Am. Mineral.* **80**, 305–318 (1995).

12. M. H. Cohen, G. S. Grest, Liquid-glass transition, a free-volume approach. *Phys. Rev. B*. **20**, 1077 (1979).

13. D. R. Neuville, P. Richet, Viscosity and mixing in molten (Ca, Mg) pyroxenes and garnets. *Geochim. Cosmochim. Acta*. **55**, 1011–1019 (1991).

14. P. Richet, Residual and configurational entropy: Quantitative checks through applications of Adam-Gibbs theory to the viscosity of silicate melts. *J. Non-Cryst. Solids*. **355**, 628–635 (2009).

15. D. E. Day, Mixed alkali glasses - Their properties and uses. *J. Non-Cryst. Solids*. **21**, 343–372 (1976).

16. T. J. Kaptchuk, Effect of interpretive bias on research evidence. *BMJ*. **326**, 1453–1455 (2003).

17. D. Giordano, J. K. Russell, Towards a structural model for the viscosity of geological melts. *Earth Planet. Sci. Lett.* **501**, 202–212 (2018).

18. C. Le Losq, D. R. Neuville, Molecular structure, configurational entropy and viscosity of silicate melts: Link through the Adam and Gibbs theory of viscous flow. *J. Non-Cryst. Solids*. **463**, 175–188 (2017).

19. G. N. Greaves, A. L. Greer, R. S. Lakes, T. Rouxel, Poisson’s ratio and modern materials. *Nat. Mater.* **10**, 823–837 (2011).

20. I. M. Hodge, Enthalpy relaxation and recovery in amorphous materials. *J. Non-Cryst. Solids*. **169**, 211–266 (1994).

21. G. Liu, D. Zhao, Y. Zuo, Modified Adam–Gibbs models based on free volume concept and their application in the enthalpy relaxation of glassy polystyrene. *J. Non-Cryst. Solids*. **417**–**418**, 52–59 (2015).

22. H. Shintani, H. Tanaka, Frustration on the way to crystallization in glass. *Nat. Phys.* **2**, 200–206 (2006).

23. W. H. Zachariasen, The atomic arrangement in glass. *J. Am. Chem. Soc.* **54**, 3841–3851 (1932).

24. G. N. Greaves, Exafs and the structure of glass. *J. Non-Cryst. Solids*. **71**, 203–217 (1985).

25. G. N. Greaves, K. L. Ngai, Reconciling ionic-transport properties with atomic structure in oxide glasses. *Phys. Rev. B*. **52**, 6358–6380 (1995).

*Supplementary Material References*

**To be done prior to submission**

**Acknowledgments:** CLL thanks Malcolm Sambridge (Seismology & Mathematical Geophysics, RSES, Australian National University), Lexing Xie and Cheng Soon Ong (CECS, Australian National University), and S. K. Lee (Seoul National University) for various discussions regarding optimization, machine learning, and the implementation of the present model. **Funding:** CLL acknowledges funding from a “Chaire d’Excellence” from the Université de Paris, from the Australian Research Council Laureate Fellowship FL1600000 to Prof. Hugh O’Neill (ANU-RSES) as well as from the Postdoctoral Fellowship of the Carnegie Institution for Science during the realisation of this project. **Author contributions:** CLL designed the study, collected the data, performed Raman and viscosity experiments, developed the computer code and drafted the manuscript. DN and BOM helped in the design of the neural network. CLL, BOM and DN performed Raman measurements. All authors contributed to the final version of the manuscript. **Competing interests:** Authors declare no competing interests. **Data and materials availability:** All the data are available in the main text or the supplementary materials. The computer code to reproduce the results of this study is available as a Python library at the web address <https://github.com/charlesll/neuravi> .

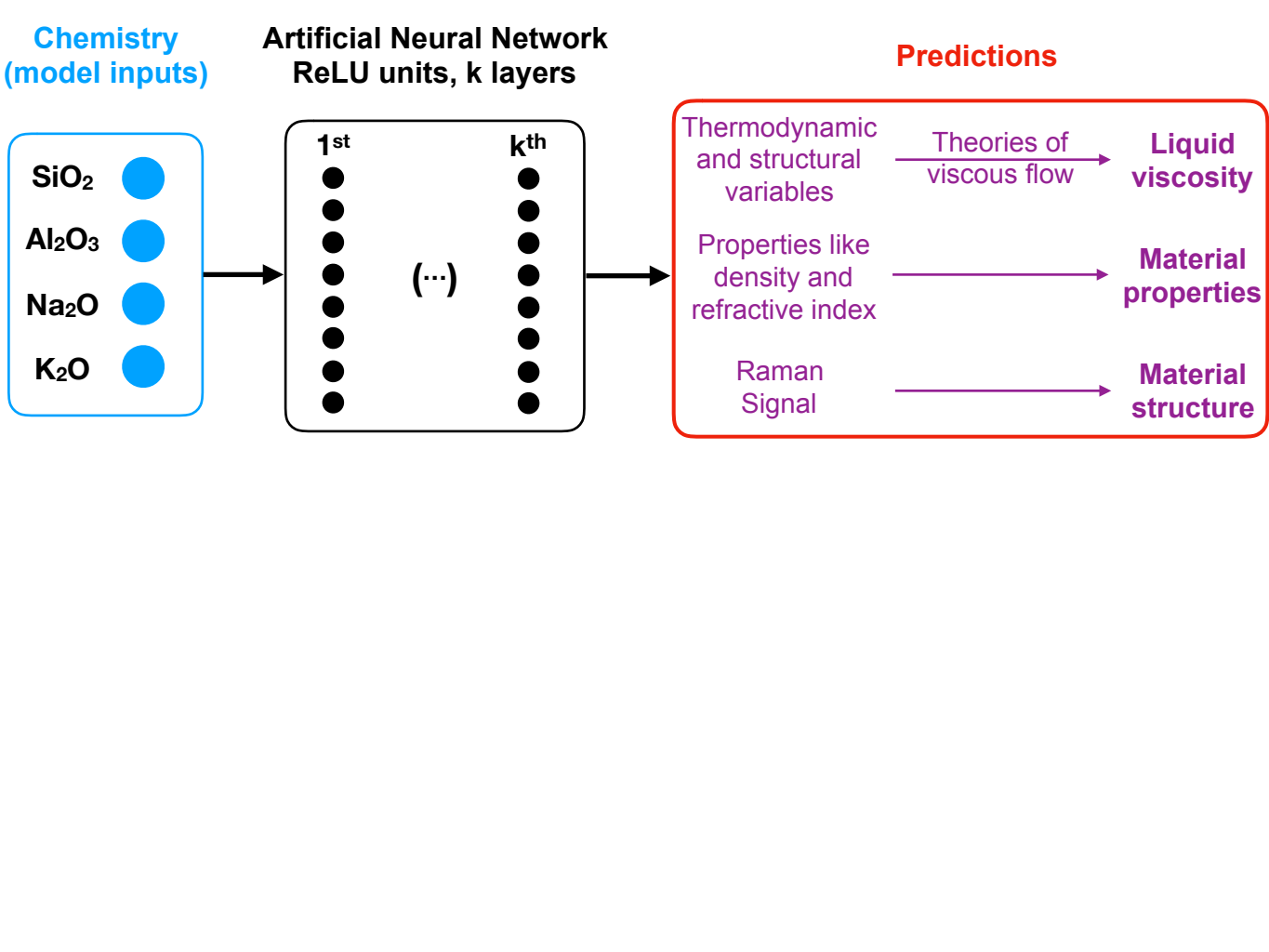
**Supplementary Materials:**

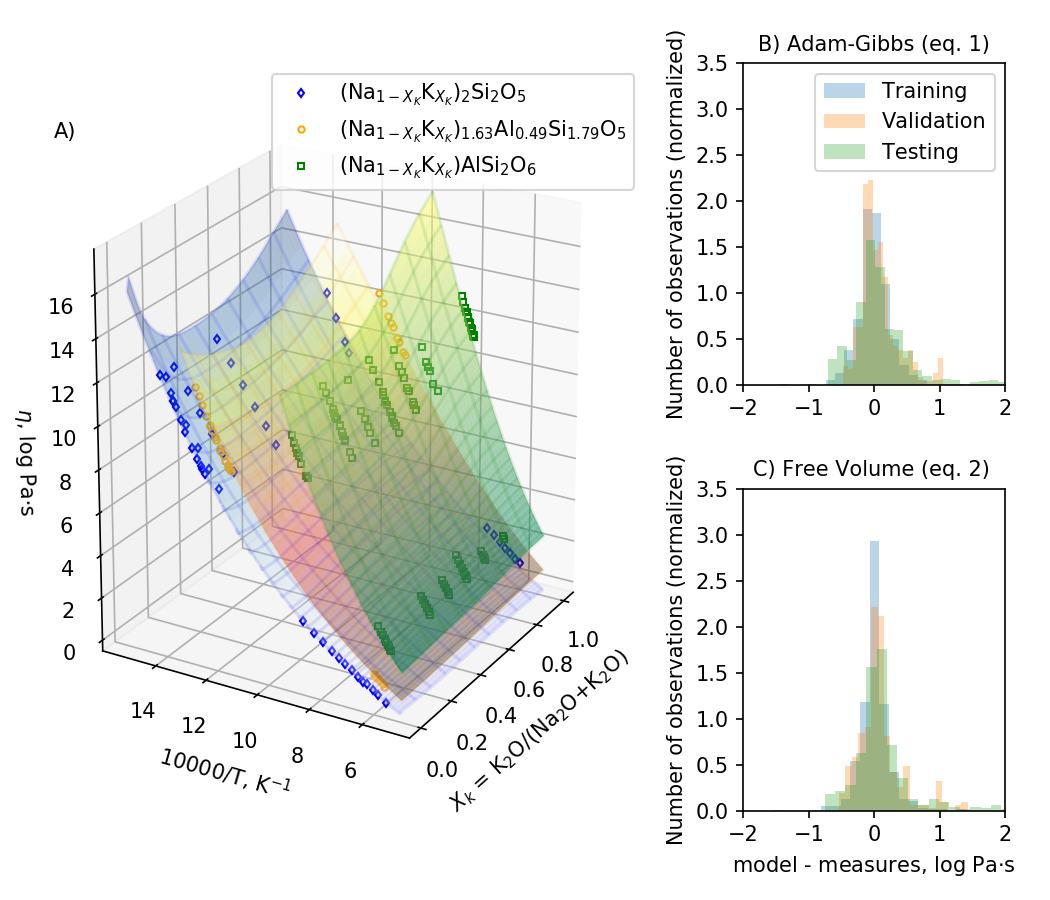
Materials and Methods

Supplementary Text

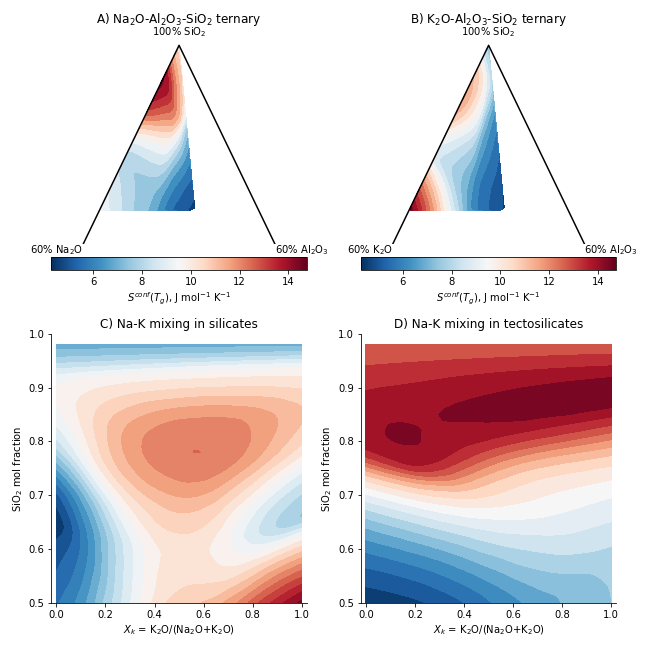
Figures S1-S7

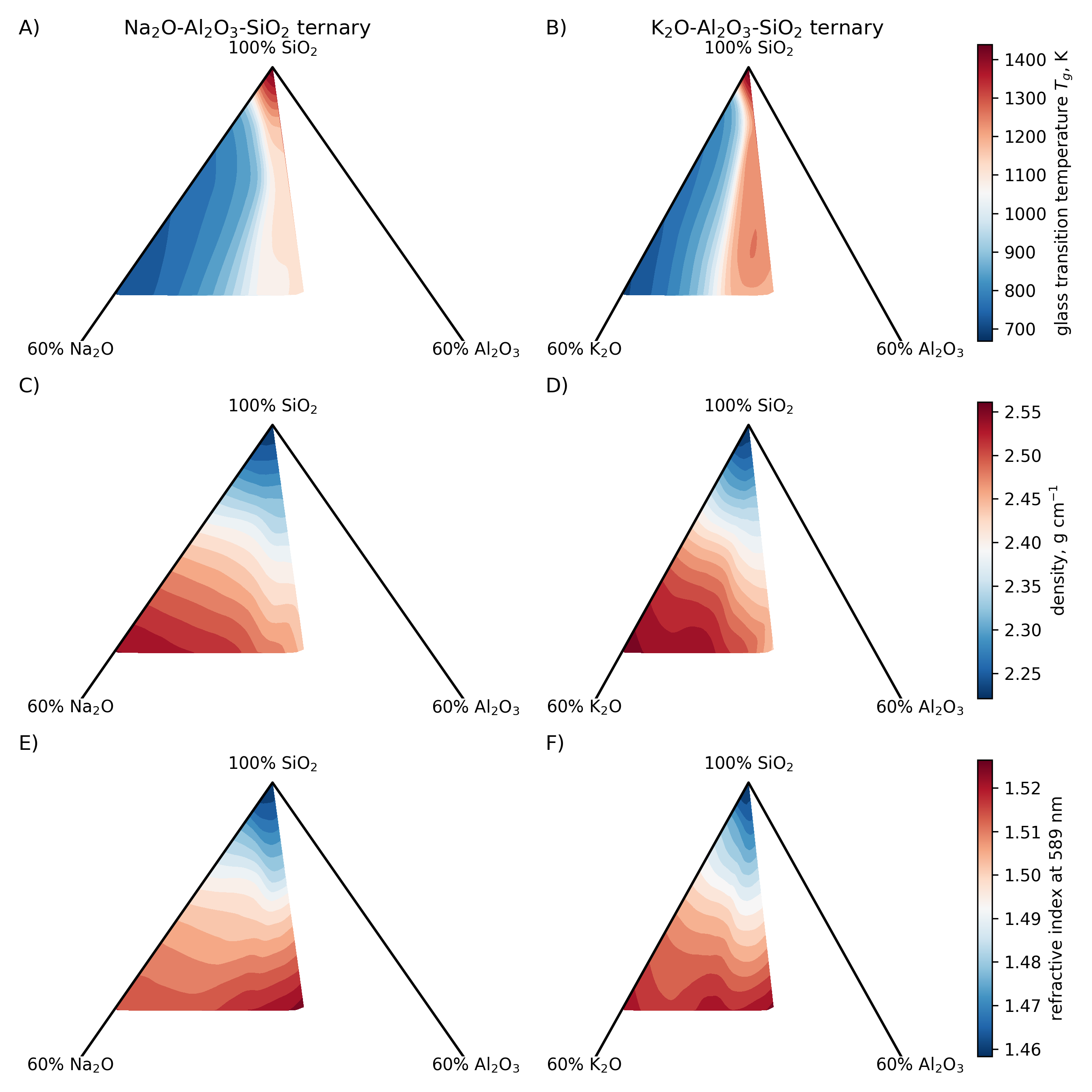
Tables S1-S3: Composition of the synthesized glasses.

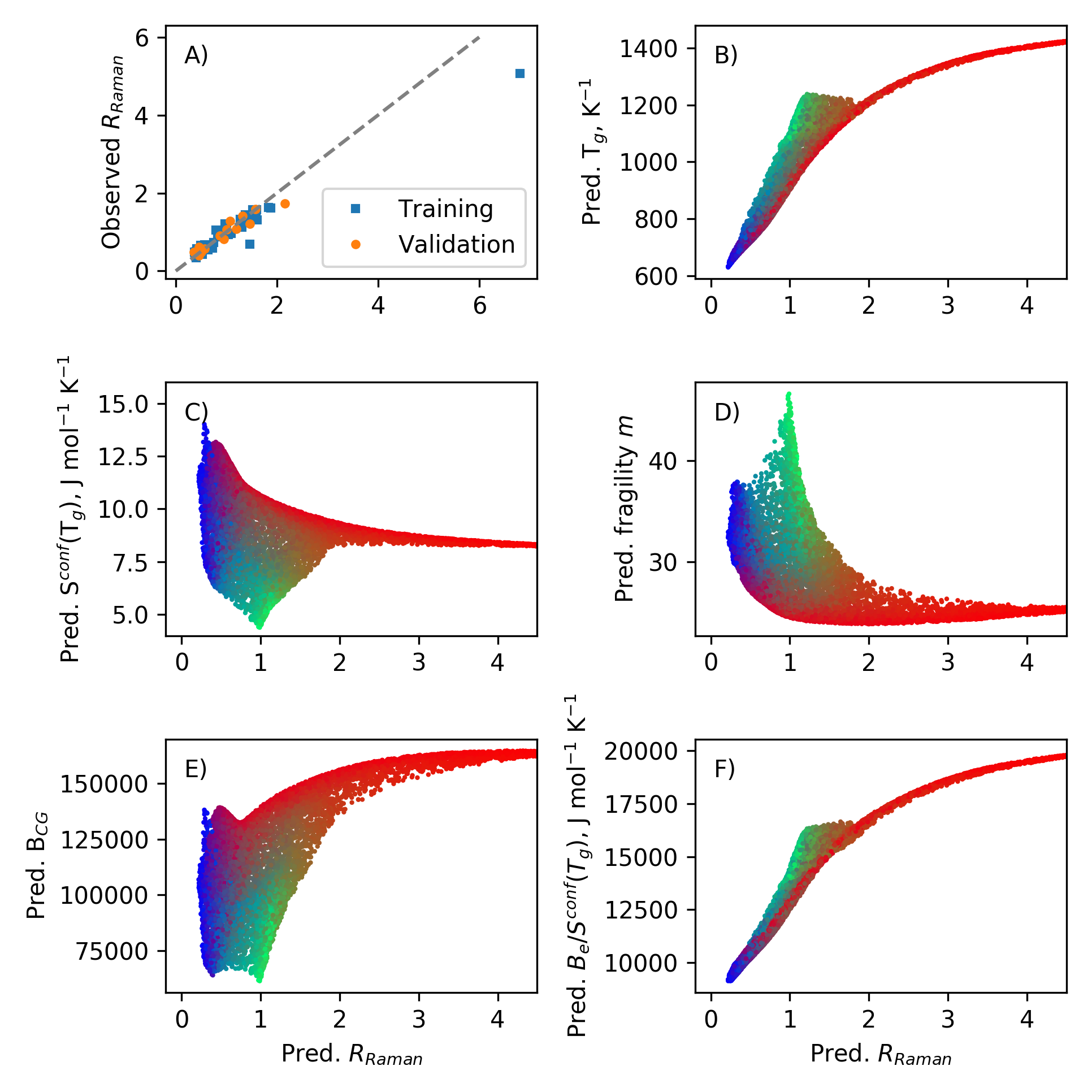
**Fig. 1.** **Presentation of the model.** The melt composition is given in input to a deep neural network that predicts various melt and glass properties as well as their Raman vibrational properties (see text for details). Viscosity can then be calculated via different theories, and the relationships between chemistry, structure and properties of the melt and glasses systematically explored using this framework.

****

**Figure 2: Example of viscosity predictions and observation of viscosity precision. (A)** Measurements (symbols) and model predictions using eq. (1) (colored wireframe) are reported as a function of inverse temperature for three different melt series with variable metal to aluminium ratios. (B) and (C) are histograms of the difference between model viscosity predictions using eqs. 1 and 2 and measurements for the different data subsets. See also Table S3 and Figure S3.

**Figure 3: Maps of configurational entropy at the glass transition temperature.** *Sconf(Tg)*, is shown in the A) sodium and B) potassium aluminosilicate systems as well as during Na-K mixing in C) silicates (no Al) and D) tectosilicates (Al/(Na+K) = 1) compositions. Addition of aluminum generally result in decreasing *Sconf(Tg)*, which reach minimal values close to the *nepheline-orthoclase* composition. During Na-K mixing, excess entropy appears when those elements are network modifiers (in silicate, C) but is not detected in Al-rich compositions (e.g. in aluminosilicates, D).

**Figure 4: Maps of glass transition temperature, glass density and glass refractive index**. Glass transition temperature varies strongly with the (Na+K)/Al ratio and silica content in both the sodium (A) and potassium (B) aluminosilicate systems, while glass density (C, D) or optical refractive index depend mostly on the glass silica content (E, F).

**Figure 5: Relationship between glass Raman spectra and glass properties**. The ratio of the intratetrahedral to intertetrahedral vibrations, *RRaman* (see text for explanation), is a proxy for network topology. It is well reproduced by the present model despite a limited training set (A); the model thus captured well how glass composition is linked to their Raman signal and thus structure. *RRaman* correlates well with *Tg* (B) or *Be/Sconf(Tg)* (F), but variations with *Sconf(Tg)*, *m* or *BCG* are of a more complex nature (C, D, E). Each point is part of 10,000 randomly generated compositions in the quaternary diagram. Symbol colors are in RGB format and embed compositional information as red is scaled mol fraction of SiO2, green is scaled mol fraction of Al2O3, and blue is scaled mol fraction of K2O+Na2O. This allows to observe a clear compositional mapping in variations of *Sconf(Tg)*, *m* or *BCG*.