

Title: Deep learning prediction of glass and melt properties**Authors :** Charles Le Losq^{1,2,3,*}, Andrew Valentine², Bjorn O. Mysen³, Daniel R. Neuville¹**Affiliations :**¹Université de Paris, Institut de physique du globe de Paris, CNRS-UMR 7154, Paris 75005,

5 France

²Research School of Earth Sciences, Australian National University, Canberra 2601, Australia³Geophysical Laboratory, Carnegie Institution for Science, Washington D.C. 20001, U.S.A.*Correspondence to: lelosq@ipgp.fr10 **Abstract:** (125/125 w)

Knowledge of the link between chemistry, structure and properties of molten rocks and glasses is essential for Earth sciences and the glass industry. No general model allows predicting and studying this link. Such a model is needed for practical application of experimentally-derived properties, often measured in chemically simple systems, to geological and industrial processes.

15 Here, a deep learning framework embedding thermodynamic equations predicts melts and glasses properties, including viscosity, density, and Raman signals. Trained on alkali aluminosilicate compositions, it quantifies the link between nano-scale changes in lava composition/structure and the dynamic of mega-scale eruptions of silicic volcanoes, like Yellowstone (U.S.A.). This is a glimpse into the possibilities offered by this approach, which 20 provides a new way to build models of material properties for various applications.

One sentence summary (112/125 characters): Deep learning bridges the gap between data and models for practical material property prediction and exploration

Main Text:

25 How do molten silicates move? How do they exchange heat with other media? How do they crystallize? These fundamental questions underpin many practical problems, including the dynamics of volcanic eruptions (*1*), the formation of rocks (*2*), and the manufacturing of novel glass, glass-ceramic and ceramic materials. Addressing them requires knowledge of physical properties such as viscosity, heat capacity and entropy, which are ultimately governed by the
30 liquid's composition and associated atomic/ionic structure. At present, this information is inferred from specific experimental observations and/or thermodynamic or molecular dynamics (MD) simulations (e.g. *3, 4*); these are usually restricted to a small range of liquid chemical compositions. Thermodynamic or empirical models may provide an intermediate way of predicting macroscopic properties, but these are usually highly simplified, and limited to
35 predicting a handful of properties in a restricted compositional range (e.g., see for viscosity *5–8*). Here, we show that these shortcomings can be resolved by combining theoretical equations and deep learning to construct a general predictive framework. This allows prediction of liquid properties under a wide range of realistic circumstances, and enables investigation of the fundamental links between material composition, structure and properties.

40

As a concrete example, we consider aluminosilicate melts and glasses, materials central to the Earth and material sciences (*9, 10*). We particularly focus on compositions in the $K_2O-Na_2O-Al_2O_3-SiO_2$ system, central to understanding the activity of silicic supervolcanoes like Yellowstone (U.S.A.) or Toba (Indonesia) (*11*), and for which a fairly complete dataset is
45 available (Fig. S1). Viscosity is key to this problem (*11*), but the underlying structural and thermodynamic controls have yet to be properly characterized. A better understanding of – and

ability to predict – the properties of alkali aluminosilicate glasses is also valuable beyond volcanology. For example, it can aid the development of technological glass materials such as smartphone screens.

50

To address any such problems, we require a model that links melt composition (x) and structure to physical properties and intensive variables such as temperature (T). To build such a model, we construct a deep neural network in Pytorch, containing multiple hidden layers of rectifier activation units, that takes an input composition, and predicts directly-observable glass properties 55 including density, refractive index and Raman spectra (Fig. 1a). It also outputs the latent variables (such as configurational entropy, $S^{conf}(T_g)$, a property that reflects the melt structure) required to predict melt viscosity (η) within five theoretical and empirical models commonly used for reproducing experimentally-observed variations of viscosity with temperature for silicate melts: Adam-Gibbs, MYEGA, Avramov-Milchev, Tamman-Vogel-Fulcher and Free 60 Volume Theory (Supplementary Material). While those models are popular for silicate melts, no strong consensus towards any one model appropriate for all liquids exists. To circumvent such problem, we propose a “trans-theoretical” approach, whereby the neural network simultaneously enables predictions across multiple theories, and is trained to provide best average performance across the selected theories. The neural network also provides optimal common values for 65 parameters that appear in multiple theories, such as the glass transition temperature T_g . To generate the deep learning framework used for predictions, first, 2000 neural networks containing various randomly selected numbers of layers and activation units per layers were trained with the ADAM optimizer. Generalization was promoted by the use of dropout, early stopping and the multi-task learning strategy (12–14). Final predictions are made via averaging

70 the predictions of the best 10 trained neural networks (15). Details are provided in Supplementary Materials.

After training, the deep learning framework allows accurate predictions of liquid and glass properties, including structure-dependent properties such as Raman spectra. Trans-theoretical predictions of η (Fig. 1b, Figure S3) are possible with good precision ($\sigma_\eta < 0.4 \log \text{Pa}\cdot\text{s}$ on 75 unseen data, Table S3; for comparison, $\sigma_\eta > 0.6 \log \text{Pa}\cdot\text{s}$ for the best empirical magma viscosity models (e.g. 5)). Known viscous T_g and $S^{conf}(T_g)$ are predicted within 16 K and $0.8 \text{ J mol}^{-1} \text{ K}^{-1}$, respectively, and glass density and refractive index are predicted within 0.009 g cm^{-3} and 0.005, respectively (Fig. S4). Despite a limited dataset (Fig. S1), global variations of Raman signals 80 have been captured well (Fig. 1c, S4) and can be predicted with a reasonable error (Supplementary Text). The deep learning framework thus embeds structural information, and allows estimation of structural parameters from Raman spectra including the ratio of intra- and inter-tetrahedral aluminosilicate vibrations, R_{Raman} (Fig. 1c). This serves as a proxy for glass network topology (the 3D connectivity of the polyhedral $\text{SiO}_2\text{-AlO}_2$ network), and is linked to 85 variations in melt properties (16).

With the deep learning framework, we can explore why viscosity varies substantially according to the concentration of Al and K in $\text{Na}_2\text{O}\text{-K}_2\text{O}\text{-Al}_2\text{O}_3\text{-SiO}_2$ melts. We observe the known division between topological and chemical effects (17) on different properties. Glass network topology, 90 broadly speaking, correlates with the glass transition temperature (Fig. 2a), and with quantities proportional to energy barriers opposed to ionic mobility (Fig. 2b), or the glass refractive index (Supplementary Materials). The glass transition temperature exhibits detailed compositional

dependence, in agreement with the well-known mixed alkali effect (MAE, 18) resulting from metal cation interactions within the SiO_4^{4-} polyhedral network (e.g. 7). This chemical effect
95 strongly affects properties influenced by cationic interactions and steric hindrance effects, such as the configurational entropy (7, 17, 19, 20), or the term C_{FV} that encompass local cationic influences on melt free volume in the Free Volume theory (Fig. 2c,d, Supplementary Text). The deep learning framework allows such effects to be seen, and the ability to switch between theories further allows observing relationships between latent variables from different theories
100 (Supplementary Text).

The above observations highlight the known role of cationic mixing on macroscopic variables such as $S^{\text{conf}}(T_g)$. Cationic mixing can be random (17, 20) or not (21, 22), and is usually difficult to predict. The deep learning framework solves this problem by enabling systematic
105 quantification and visualization of such phenomena (Fig. 3). For the system analyzed here, increasing [Al] leads to decreasing $S^{\text{conf}}(T_g)$ (Fig. 3a,b). Furthermore, Al-to-alkali ratio largely affects the MAE: without Al, the MAE results in an entropy excess (Fig. 3c) and, hence, in large decreases in melt viscosity ($\eta \propto 1/S^{\text{conf}}(T_g)$, 19). Increasing [Al]/[Na+K] leads to changing the role of alkali metals in the network (23). Na and K segregate in different molecular nano-
110 environments (7, 24), inducing less and less excess entropy of mixing as [Al]/[Na+K] increases (Fig. 3d,e,f). In addition, K-rich and Al-rich melts show higher local order (4) resulting in small $S^{\text{conf}}(T_g)$ values (Fig. 3b,f).

Most of the lavas emitted at silicic volcanic systems, like Yellowstone or Long Valley (U.S.A.),
115 present dry chemical compositions made of more than 95 % of Na_2O , K_2O , Al_2O_3 and SiO_2 .

Melts in the Na₂O-K₂O-Al₂O₃-SiO₂ system can thus be considered as analogues of the lavas involved in silicic volcanic eruptions. With this in mind, we are now in a position to understand why eruptions of silicic volcanoes may be more explosive if the magma is rich in K and Al, as suggested by (11). Indeed, among the usual critical parameters driving the dynamic of volcanic 120 eruptions such as pre-eruptive volatile content, degassing path and nano-cristallization (11, 25, 26), explosive or effusive eruptions of silicic volcanoes respectively appear to involve lavas more or less rich in Al and K (Figure 4). From the deep learning framework predictions, this originates from a decrease in $S^{conf}(T_g)$ driven by ongoing network connectivity and nano-structuration as [Al] and [K] increase (Figs. 3, 4), as those elements respectively promote 125 network polymerisation and nano-structuration in aluminosilicates (4, 23). The chemical separation between effusive and explosive eruptions at silicic volcanic centers is associated, according to the present work, with limits in R_{Raman} and $S^{conf}(T_g)$ of ~1.5 and ~9.0 J mol⁻¹ K⁻¹, respectively (Fig. 4). Those values are derived from the analysis of simplified alkali aluminosilicates and thus may vary slightly in the case of natural rhyolite. In any case, those 130 results demonstrate that a more complete version of the present framework, embedding the effects of Ca, Mg, Fe and water, would provide the ability to explore and quantify even further the links between magma composition, structure, properties, and volcanic eruptive styles.

To conclude, the results presented here highlight how the deep learning framework can be used 135 for practical property predictions in the Earth sciences. Here, it reveals the fine structural and thermodynamic controls on magma viscosity, which, in turn, govern the dynamics of volcanic eruptions. The deep learning framework can readily be extended to include quantities of interest across a range of domains and applications, with the present results demonstrating that a

combination of neural networks and physical/thermodynamic models can offer new perspectives
140 on long-standing problems in many disciplines.

References and Notes. (30 citations max main text)

1. D. B. Dingwell, Volcanic Dilemma - Flow or Blow? *Science*. **273**, 1054–1055 (1996).
2. N. L. Bowen, *The evolution of igneous rocks* (Dover Publications, 1956).
3. 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.
4. 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).
5. D. Giordano, J. K. Russell, D. B. Dingwell, Viscosity of magmatic liquids: A model. *Earth Planet. Sci. Lett.* **271**, 123–134 (2008).
6. J. C. Mauro, A. J. Ellison, D. C. Allan, M. M. Smedskjaer, Topological Model for the Viscosity of Multicomponent Glass-Forming Liquids. *Int. J. Appl. Glass Sci.* **4**, 408–413 (2013).
7. 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).
8. K. Starodub, G. Wu, E. Yazhenskikh, M. Müller, A. Khvan, A. Kondratiev, An Avramov-based viscosity model for the SiO₂-Al₂O₃-Na₂O-K₂O system in a wide temperature range. *Ceram. Int.* **45**, 12169–12181 (2019).
9. B. O. Mysen, P. Richet, *Silicate Glasses and Melts* (Elsevier, ed. 2nd, 2019).
10. 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>).
11. D. Di Genova, S. Kolzenburg, S. Wiesmaier, E. Dallanave, D. R. Neuville, K. U. Hess, D. B. Dingwell, A compositional tipping point governing the mobilization and eruption style of rhyolitic magma. *Nature*. **552**, 235 (2017).
12. R. Caruana, Multitask Learning. *Mach. Learn.* **28**, 41–75 (1997).

13. N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, R. Salakhutdinov, Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *J. Mach. Learn. Res.* **15**, 1929–1958 (2014).
14. I. Goodfellow, Y. Bengio, A. Courville, *Deep Learning* (MIT Press, 2016; <http://www.deeplearningbook.org>).
15. L. Breiman, L. Breiman, in *Machine Learning* (1996), pp. 123–140.
16. D. Giordano, J. K. Russell, Towards a structural model for the viscosity of geological melts. *Earth Planet. Sci. Lett.* **501**, 202–212 (2018).
17. D. R. Neuville, P. Richet, Viscosity and mixing in molten (Ca, Mg) pyroxenes and garnets. *Geochim. Cosmochim. Acta* **55**, 1011–1019 (1991).
18. J. O. Isard, The mixed alkali effect in glass. *J. Non-Cryst. Solids* **1**, 235–261 (1969).
19. P. Richet, Viscosity and configurational entropy of silicate melts. *Geochim. Cosmochim. Acta* **48**, 471–483 (1984).
20. D. R. Neuville, B. O. Mysen, Role of aluminium in the silicate network: In situ, high-temperature study of glasses and melts on the join SiO₂-NaAlO₂. *Geochim. Cosmochim. Acta* **60**, 1727–1737 (1996).
21. D. R. Neuville, Viscosity, structure and mixing in (Ca, Na) silicate melts. *Chem. Geol.* **229**, 28–41 (2006).
22. S. K. Lee, Structure and the extent of disorder in quaternary (Ca-Mg and Ca-Na) aluminosilicate glasses and melts. *Am. Mineral.* **90**, 1393–1401 (2005).
23. C. Le Losq, D. R. Neuville, P. Florian, G. S. Henderson, D. Massiot, The role of Al³⁺ on rheology and structural changes of sodium silicate and aluminosilicate glasses and melts. *Geochim. Cosmochim. Acta* **126**, 495–517 (2014).
24. C. Le Losq, D. R. Neuville, Effect of the Na/K mixing on the structure and the rheology of tectosilicate silica-rich melts. *Chem. Geol.* **346**, 57–71 (2013).
25. J. Andújar, B. Scaillet, Relationships between pre-eruptive conditions and eruptive styles of phonolite–trachyte magmas. *Lithos* **152**, 122–131 (2012).
26. F. Cáceres, F. B. Wadsworth, B. Scheu, M. Colombier, C. Madonna, C. Cimarelli, K.-U. Hess, M. Kaliwoda, B. Ruthensteiner, D. B. Dingwell, Can nanolites enhance eruption explosivity? *Geology* (2020), doi:10.1130/G47317.1.

Supplementary Material References

27. O. V. Mazurin, M. V. Streltsina, Shvaiko-Shvaikovskaya, *Handbook of glass data* (Elsevier, 1987), vol. C and B.

28. C. Le Losq, A. J. Berry, M. A. Kendrick, D. R. Neuville, H. St. C. O'Neill, Determination of the oxidation state of iron in Mid-Ocean Ridge basalt glasses by Raman spectroscopy. *Am. Mineral.* **104**, 1032–1042 (2019).
29. C. M. Bishop, *Pattern recognition and machine learning* (Springer, New York, 2006), *Information science and statistics*.
30. G. Adam, J. H. Gibbs, On the temperature dependence of cooperative relaxation properties in glass-forming liquids. *J. Chem. Phys.* **43**, 139–146 (1965).
31. M. H. Cohen, G. S. Grest, Liquid-glass transition, a free-volume approach. *Phys. Rev. B.* **20**, 1077 (1979).
32. I. Avramov, A. Milchev, Effect of disorder on diffusion and viscosity in condensed systems. *J. Non-Cryst. Solids.* **104**, 253–260 (1988).
33. J. C. Mauro, Y. Yue, A. J. Ellison, P. K. Gupta, D. C. Allan, Viscosity of glass-forming liquids. *Proc. Natl. Acad. Sci.* **106**, 19780–19784 (2009).
34. X. Glorot, A. Bordes, Y. Bengio, in *International Conference on Artificial Intelligence and Statistics* (2011), pp. 315–323.
35. J. Bergstra, Y. Bengio, Random search for hyper-parameter optimization. *J. Mach. Learn. Res.* **13**, 281–305 (2012).
36. G. N. Greaves, A. Fontaine, P. Lagarde, D. Raoux, S. J. Gurman, Local structure of silicate glasses. *Nature*. **293**, 611–616 (1981).
37. G. N. Greaves, EXAFS and the structure of glass. *J. Non-Cryst. Solids.* **71**, 203–217 (1985).
38. G. N. Greaves, EXAFS, glass structure and diffusion. *Philos. Mag. Part B.* **60**, 793–800 (1989).
39. A. Meyer, J. Horbach, W. Kob, F. Kargl, H. Schober, Channel formation and intermediate range order in sodium silicate melts and glasses. *Phys. Rev. Lett.* **93**, 1–4 (2004).
40. F. Kargl, A. Meyer, M. M. Koza, H. Schober, Formation of channels for fast-ion diffusion in alkali silicate melts: A quasielastic neutron scattering study. *Phys. Rev. B.* **74**, 014304 (2006).
41. G. N. Greaves, K. L. Ngai, Reconciling ionic-transport properties with atomic structure in oxide glasses. *Phys. Rev. B.* **52**, 6358–6380 (1995).
42. J. Stebbins, NMR evidence for five-coordinated silicon in a silicate glass at atmospheric pressure. *Nature*. **351**, 638–639 (1991).

43. P. Y. Huang, S. Kurasch, A. Srivastava, V. Skakalova, J. Kotakoski, A. V. Krasheninnikov, R. Hovden, Q. Mao, J. C. Meyer, J. Smet, D. A. Muller, U. Kaiser, Direct Imaging of a Two-Dimensional Silica Glass on Graphene. *Nano Lett.* **12**, 1081–1086 (2012).
44. P. Y. Huang, S. Kurasch, J. S. Alden, A. Shekhawat, A. A. Alemi, P. L. McEuen, J. P. Sethna, U. Kaiser, D. A. Muller, Imaging Atomic Rearrangements in Two-Dimensional Silica Glass: Watching Silica’s Dance. *Science*. **342**, 224–227 (2013).
45. B. O. Mysen, Experimental, in situ, high-temperature studies of properties and structure of silicate melts relevant to magmatic processes. *Eur. J. Mineral.* **7**, 745–766 (1995).
46. B. O. Mysen, A. Lucier, G. D. Cody, The structural behavior of Al^{3+} in peralkaline melts and glasses in the system $\text{Na}_2\text{O}-\text{Al}_2\text{O}_3-\text{SiO}_2$. *Am. Mineral.* **88**, 1668–1678 (2003).
47. H. Hui, Y. Zhang, Toward a general viscosity equation for natural anhydrous and hydrous silicate melts. *Geochim. Cosmochim. Acta*. **71**, 403–416 (2007).
48. X. Duan, A model for calculating the viscosity of natural iron-bearing silicate melts over a wide range of temperatures, pressures, oxygen fugacities, and compositions. *Am. Mineral.* **99**, 2378–2388 (2014).
49. S. L. L. Webb, Configurational heat capacity of $\text{Na}_2\text{O}-\text{CaO}-\text{Al}_2\text{O}_3-\text{SiO}_2$ melts. *Chem. Geol.* **256**, 92–101 (2008).
50. J. K. Russell, D. Giordano, Modelling configurational entropy of silicate melts. *Chem. Geol.* **461**, 140–151 (2017).
51. P. Richet, Y. Bottinga, Heat capacity of aluminum-free liquid silicates. *Geochim. Cosmochim. Acta*. **49**, 471–486 (1985).
52. P. Richet, Y. Bottinga, Glass transitions and thermodynamic properties of amorphous SiO_2 , $\text{NaAlSi}_n\text{O}_{2n+2}$ and KAlSi_3O_8 . *Geochim. Cosmochim. Acta*. **48**, 453–470 (1984).
53. P. Richet, R. A. Robie, J. Rogez, B. S. Hemingway, P. Courtial, C. Téqui, Thermodynamics of open networks: ordering and entropy in NaAlSiO_4 glass, liquid and polymorphs. *Phys. Chem. Miner.* **17**, 385–394 (1990).
54. G. N. Greaves, A. L. Greer, R. S. Lakes, T. Rouxel, Poisson’s ratio and modern materials. *Nat. Mater.* **10**, 823–837 (2011).
55. I. M. Hodge, Enthalpy relaxation and recovery in amorphous materials. *J. Non-Cryst. Solids*. **169**, 211–266 (1994).

56. 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).

145 **Acknowledgments:** CLL thanks Malcolm Sambridge (Seismology & Mathematical Geophysics, RSES, Australian National University), Lexing Xie and Cheng Soon Ong (CECS, Australian National University), and Sung Keun Lee (Seoul National University) for various discussions regarding optimization, machine learning, and melt and glass properties.

150 **Funding:** CLL acknowledges funding from a Chaire d'Excellence IDEX19C627X/FD070/D110 from the IdEX Université de Paris ANR-18-IDEX-0001, 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 realization of this project. APV acknowledges support from the ARC via grants DE180100040 and DP200100053.

155

Author contributions: CLL designed the study, collected the data, performed Raman and viscosity experiments, and developed the deep learning framework and the associated computer code. AV and BOM helped in the design of the neural network. CLL, BOM and DN performed Raman measurements. CLL and AV drafted the manuscript. All authors contributed to the final 160 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 165 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> (open access will be provided upon acceptance).

Supplementary Materials:

170 Materials and Methods

Supplementary Text

Figures S1-S7

Tables S1-S3

175 **Fig. 1. Schematic of the deep learning framework, and prediction examples.** (a) A neural network takes input melt composition, and outputs various melt and glass properties. Once trained, relationships between chemistry, structure and properties of melts and glasses can be systematically explored. (b) Melt viscosity can be derived from these outputs using various theories with a great accuracy, as shown by examples highlighting the good match between measurements (symbols) and model (curves) predictions from the Adam-Gibbs and Free Volume theories. (c) Structural investigation are made through Raman spectra predictions, which compare well with experimental data for albite and sodium trisilicate glasses for instance. R_{Raman} , the ratio of intertetrahedral (below 670 cm^{-1}) to intratetrahedral (above 870 cm^{-1}) network vibrations, can be calculated and used for structural investigations. It increases as the connectivity and organization of the polyhedral $\text{SiO}_2\text{-Al}_2\text{O}_3$ network develops in the glass. It is thus very sensitive to 3D glass network connectivity and rigidity.

180

185

190 **Figure 2: Melt and glass properties vary in a complex way with glass network topology.** (a) The model allows observing that parameters such as the viscous glass transition temperature T_g (a) and B_{CG} , a activation energy term in the Free Volume theory (Supplementary Materials) (b), strongly vary depending on R_{Raman} . Other terms also show more complex variations, influenced by cationic mixing interactions and steric hindrance effects, such as the glass configurational entropy $S^{conf}(T_g)$ (c) or the free volume parameter C_{CG} (d). Each symbol represents the calculation for a randomly generated composition in the glass-forming domain of studied system.

195

Figure 3: Variations of the configurational entropy at T_g in $\text{K}_2\text{O-Na}_2\text{O-Al}_2\text{O}_3\text{-SiO}_2$ melts.

$S^{conf}(T_g)$ vary non-linearly with oxide contents in the ternary diagrams $\text{Na}_2\text{O-Al}_2\text{O}_3\text{-SiO}_2$ and $\text{K}_2\text{O-Al}_2\text{O}_3\text{-SiO}_2$ in (a) and (b). In silicate melts (c), a mixed alkali effect (MAE) is observed upon Na-K mixing. It disappears as $[\text{Al}]/[\text{Na+K}]$ increases (d, e, f). Al enrichment generally lead 200 to decreasing $S^{conf}(T_g)$, and minimizing the MAE. Al- and K-rich melts thus present low $S^{conf}(T_g)$.

Figure 4: Chemical, structural and entropic mapping of rhyolite eruptions. The rheological aptaitic index is calculated as $(\text{Na}_2\text{O} + \text{K}_2\text{O} + \text{CaO} + \text{MgO} + \text{FeO})/(\text{Al}_2\text{O}_3 + \text{Fe}_2\text{O}_3)$. From (11), Al-rich and K-rich rhyolite magmas, like those of Yellowstone (U.S.A.) are generally associated

with explosive eruptions. The rheological tipping point proposed by (11) find roots in how
205 increasing Al and K contents promote 3D network connectivity (a) and, thus reduces melt
entropy (b) down to a particular threshold here identified to be $\sim 9 \text{ J mol}^{-1} \text{ K}^{-1}$.

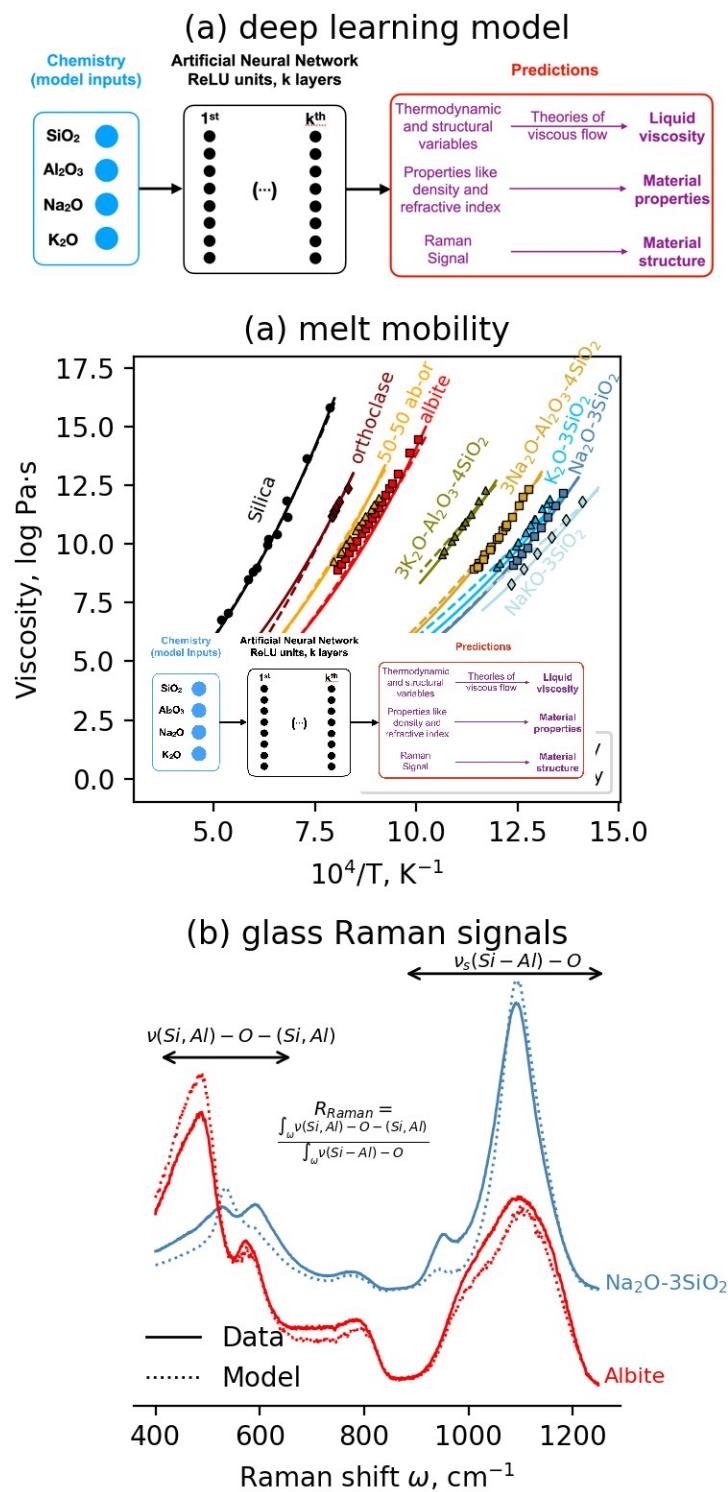
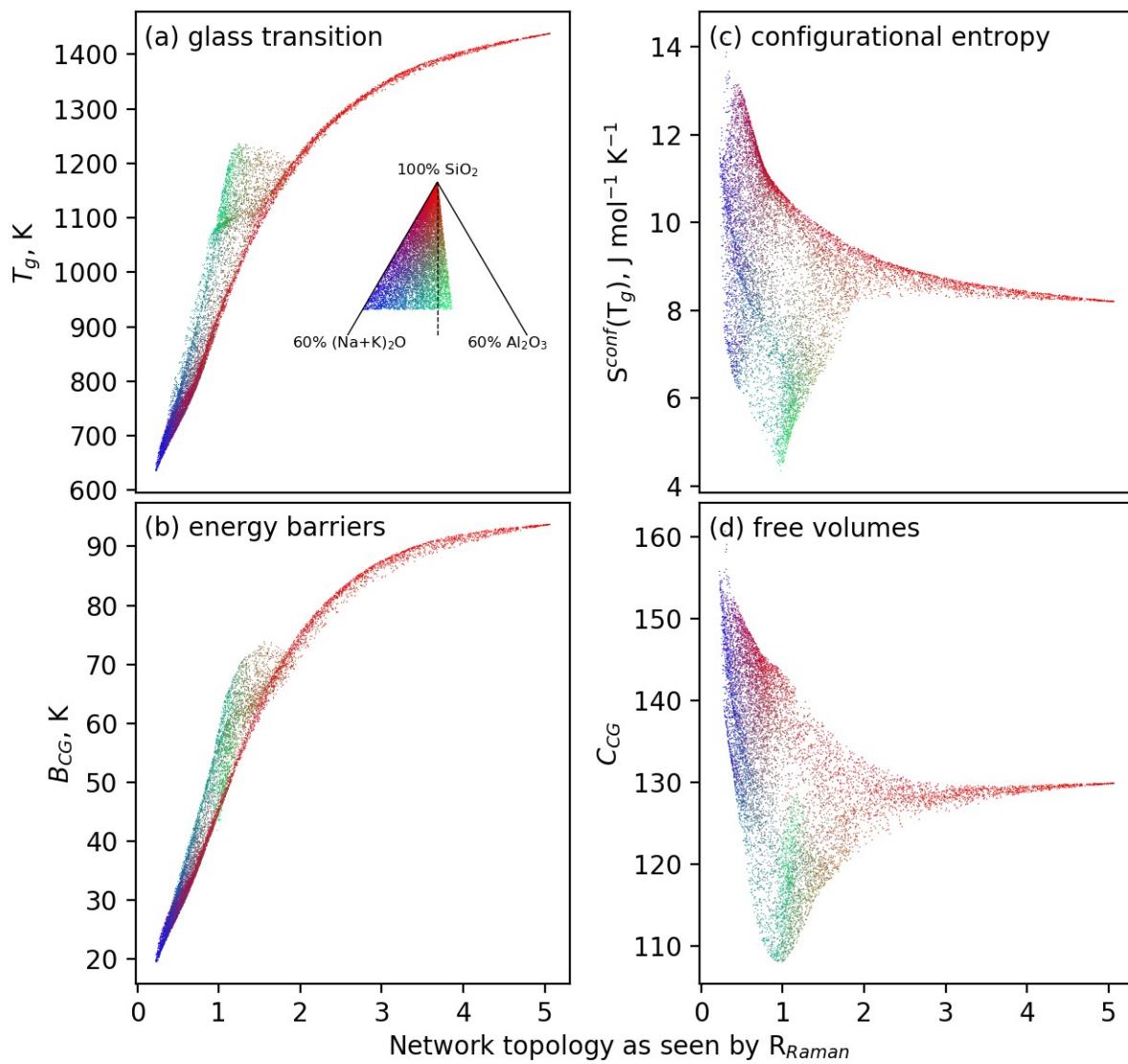
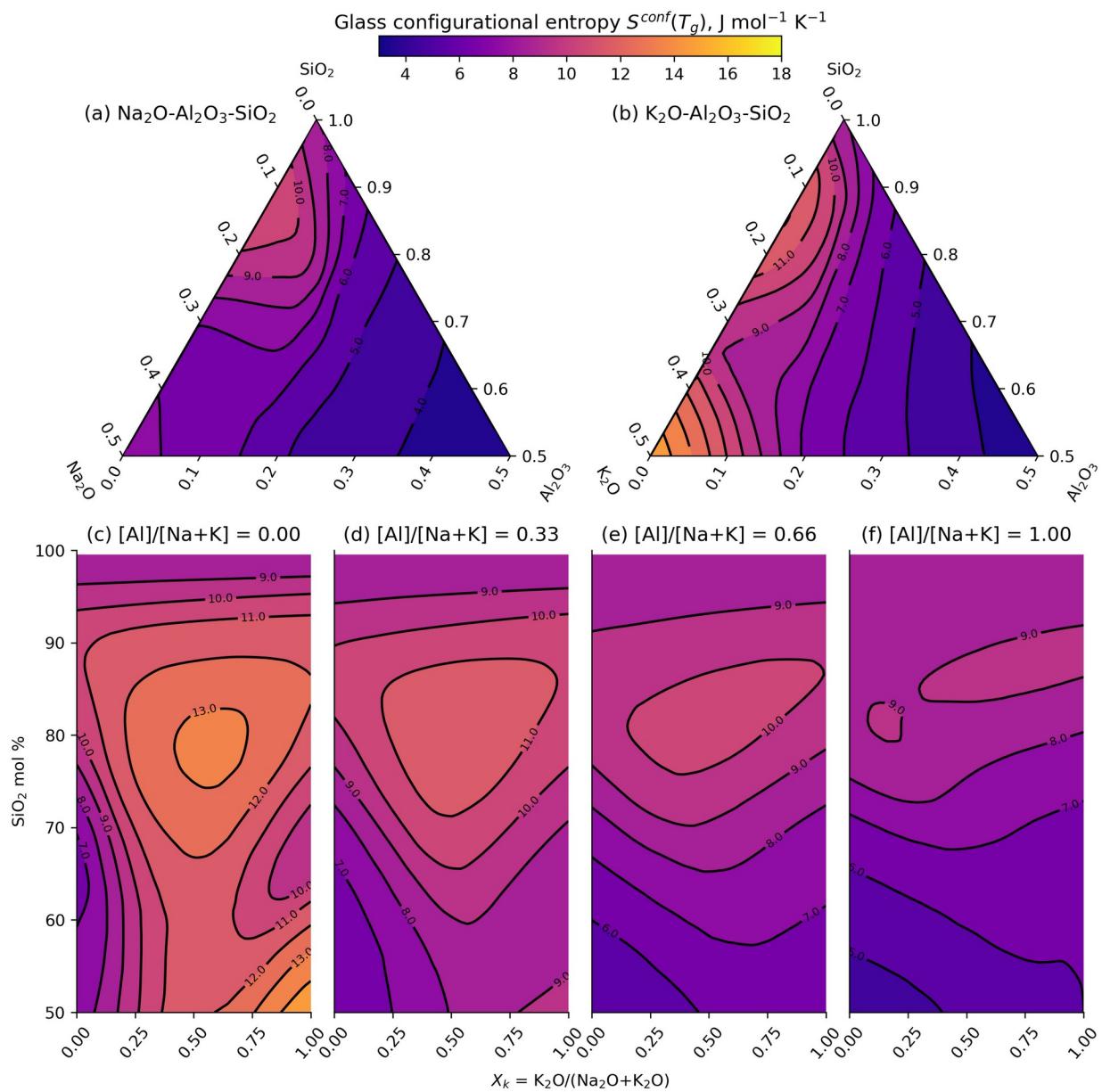


Figure 1

**Figure 2**

**Figure 3**

215

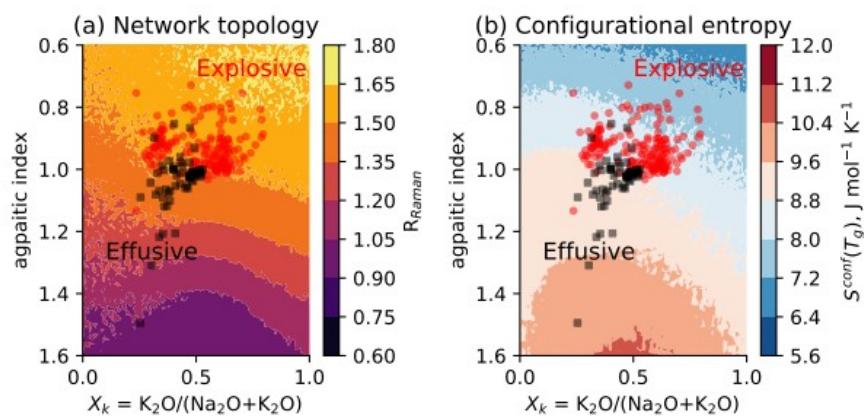


Figure 4



Supplementary Materials for

Deep learning prediction of material properties

Charles Le Losq, Andrew Valentine, Bjorn O. Mysen, Daniel R. Neuville

Correspondence to: lelosq@ipgp.fr

This PDF file includes:

Materials and Methods

Supplementary Text

Figs. S1 to S7

Tables S1 to S3

Other Supplementary Materials for this manuscript include the following:

Software repository at github.com/charlesll/neuravi

Materials and Methods

Experimental Design

Developing the model required the collection and compilation of viscosity, density, refractive index data, and Raman spectra for glasses and melts in the K₂O-Na₂O-Al₂O₃-SiO₂ quaternary diagram (Fig. S1). The viscosity of supercooled melts for peralkaline compositions in this diagram remains poorly-understood, and we conducted additional experiments to complement the existing dataset. We further compiled existing data as specified below, prior to developing the deep learning framework in the Python programming language using the PyTorch machine learning library. The codes can be run using Jupyter Notebooks. All codes and data necessary to reproduce this study can be found on the software repository Github at the web address <https://github.com/charlesll/neuravi>.

Data compilation

Existing Raman spectra and observations of optical refractive index, density and viscosity of alkali aluminosilicate glasses were selected by hand via a review of the existing literature. Cross-verification of the accuracy of viscosity data from different studies is critical and was checked on compositions including sodium trisilicate, albite and jadeite. Publications presenting deviations higher than 0.1 log Pa · s compared to the general trend on such compositions were entirely discarded. Density and refractive index come from various publications reviewed in (27). Raman spectra are published data from the IPGP and Carnegie Institution for Science

laboratories (see below for details). All data references are provided in the database available at the address https://github.com/charlesll/neuravi/data/Database_IPGP.xlsx

Sample synthesis and new measurements

To extend the scarce viscosity dataset for peralkaline aluminosilicate melts, new compositions were synthesized at IPGP in Paris from reagent-grade K_2CO_3 , Na_2CO_3 , Al_2O_3 and SiO_2 dried oxide powders, following the protocol described in (24). Viscosity and density measurements follow the protocol described in (23, 24), and the reader should consult those publications for further information. Chemical compositions have been measured using a Cameca SX50 electron microprobe (Table S1), with a 30 nA current, $U = 30$ kV, and 5 seconds of counting. Beam-induced alkali loss was minimized by working with a defocused beam that was moved continuously during the analysis. The values reported in Table S1 are the statistical mean of 10-20 individual measurements. The corresponding viscosity measurements are provided in Table S2.

Raman spectroscopy

Raman spectra of silicate and aluminosilicate glasses acquired at IPGP in Paris were recorded using a T64000 Jobin-Yvon® Raman spectrometer equipped with a confocal system, a 1024 charge-couple detector (CCD) cooled by liquid nitrogen and an Olympus® microscope. The optimal spatial resolution allowed by the confocal system is $1\text{-}2 \mu m^2$ with a $\times 100$ Olympus® objective, and the spectral resolution is 0.7 cm^{-1} . A Coherent® laser 70-C5 Ar^+ , having a wavelength of 488.1 or 514.532 nm, has been used as the excitation line. Unpolarized Raman

spectra were acquired between 20 and 1500 cm⁻¹ on pieces of glass from the starting materials that were excited with a laser power of 100-150 mW on the sample.

Further Raman spectra acquired at the Geophysical Laboratory on glasses along the K₂Si₄O₉-K₂(KAl)₄O₉ and K₂Si₄O₉-K₂(KAl)₄O₉ joins were added to the database. Those spectra were acquired with a Dilor XY confocal microRaman spectrometer equipped with a cryogenic Thompson Model 400O CCD. The 488 nm line of a SpectraPhysics model 2025 Ar⁺ laser operating at several hundred mW at the sample was used for sample excitation.

Preprocessing of the spectra was kept to minimum: (i) a linear baseline was fitted to the minima in the 700-800 and 1200-1300 cm⁻¹ portions of the spectra and then subtracted to obtain baseline-corrected spectra, (ii) the spectra were then corrected from temperature and excitation line effects (23), and (iii) the spectra were normalised to their maximum intensity such that the intensity in each spectrum varies between 0 and 1 (known as “min-max scaling” in the data science literature). Only signals in the 400-1250 cm⁻¹ range were retained as different spectra had different starting and ending Raman shift values. After pre-processing, spectra were saved in a HDF5 file for their future use.

Machine learning modelling

Datasets

Data from the above sources were compiled into a single Excel spreadsheet (see supplementary archive as well as github.com/charlesll/neuravi) that serves as a database for machine learning. Four different streams of data are present:

- $D_{viscosity}$, the database of viscosity measurements, composed of $X_{viscosity}$ chemical composition entries (mole fractions) as well as their associated temperatures (Kelvin) and $y_{viscosity}$ observations (Pa·s);
- $D_{density}$, the database of density measurements, composed of $X_{density}$ chemical composition entries (mole fractions) and $y_{density}$ observations (g cm⁻³);
- D_{Raman} , the database of Raman spectra, composed of X_{Raman} chemical composition entries (mole fractions) and y_{Raman} spectra observations (min-max scaled Raman intensities);
- $D_{optical}$, the database of optical refractive index, composed of $X_{optical}$ chemical composition entries (mole fractions) as well as their associated wavelength (μm) and $y_{refractive\ index}$ observations.

The goal of using four different set of observations was to leverage the fact that neural networks learning multiple related tasks tend to show better prediction capacities compared to those trained to perform only one task (3). $D_{viscosity}$, $D_{density}$ and $D_{optical}$ cover an important part of the glass-forming domain of alkali aluminosilicates (Fig. S1); they were thus used to train the artificial neural network with a performance oriented mindset. D_{Raman} covers a more limited set of compositions (Fig. S1). It was used as a way of improving multitask learning as well as a way of introducing structural information in the deep learning framework.

Data train-test splitting and standardisation

Machine learning algorithms are powerful interpolators, and it is important to avoid overfitting during training. This manifests itself as excellent performance on the dataset used for training, but poor-quality predictions when the algorithm is exposed to new, unseen data. This indicates that the algorithm has learned to encode the input-output relationships within the training data without capturing the underlying mathematical relationships. For artificial neural networks, overfitting is usually monitored by splitting the available dataset in three different, randomly chosen *training*, *validation* and *testing* data subsets (Fig. S1). The *training* subset is used for training the neural network while the *validation* subset is used for monitoring overfit during the training process. This enables tuning of hyperparameters to avoid this overfit (i.e. parameters of the algorithm that control the architecture of the network, the learning rate, etc.), and the selection of potential candidates for a final training step. The final predictive capacity of the trained neural networks are then evaluated using the unseen *testing* data subset. In the present study, the data were randomly separated by composition (28) to avoid the pitfall of having the same composition in the different *training*, *validation* and *testing* subsets (a phenomenon known as ‘data leakage’). $D_{viscosity}$, $D_{optical}$ and $D_{density}$ were each separated in three splits following this rule (Fig. S1). D_{Raman} was divided in only two *train* and *validation* subsets due to its small size. This is not problematic, as we do not aim at precise predictions of Raman spectra but rather use this dataset as a way to improve the predictive capacity of the trained neural network and to introduce structural knowledge.

After train-validation-test splitting, an important step in any machine learning data preprocessing is standardization of the data. In practice, appropriate data scaling is often essential to obtaining good convergence within algorithms (14). The goal of re-scaling is to promote feature variations

close to unity and ensure that all features have comparable numerical ranges: failure to do so tends to lead to instabilities in the gradient back-propagation process that is central to training neural networks. In the present study, we have implemented a custom approach. All chemical compositions inputs are in mole fractions, which corresponds to a modification of min-max (0-1) scaling. Raman spectra were normalised to be comprised between 0 and 1. Viscosity, density and refractive index were not scaled, as scaling the outputs was not found to affect network convergence. However, as we will see, when outputs are unscaled it is essential to properly initialise the bias of the output layer of the neural network to match the expected range of the predictions to be made, as done for instance for Mixture Density Network (29). After pre-processing, the different subsets were saved in HDF5 files for their future use.

Viscosity equations and the trans-theoretical nature of the deep learning framework

No consensus exists regarding a best model to explain and reproduce the variations of liquid viscosity with parameters like temperature, pressure and composition. Despite this, some models may be preferred. For instance, the Adam-Gibbs model (30) has been particularly successful in reproducing viscosity data of silicate melts (19). It assumes that liquid movements occur through cooperative molecular re-arrangements, and represents viscosity (η) as a function of temperature (T) and composition (x) via

$$\log \eta(T, x) = A_e(x) + \frac{B_e(x)}{T \left(S^{conf}(T_g, x) + \int_{T_g}^T C_p^{conf}(x) / T dT \right)}, \quad (SI)$$

with A_e representing a high-temperature limit, B_e a term proportional to the energy barriers opposed to molecular re-arrangements, and S^{conf} and C_p^{conf} the melt configurational entropy and heat capacity, respectively. T_g is the ‘glass transition temperature’, at which melt is frozen-in into

glass upon quench (for convenience, we adopt hereafter the empirical definition of T_g as equal to the temperature for which $\eta = 10^{12}$ Pa·s).

Alternatively, one might adopt the Free Volume model (31), which states that melts present liquid-like and solid-like molecular cells, their mobility being ensured by atomic diffusivity within/between liquid-like cells. This takes the form

$$\log\eta(T, x) = A_{FV}(x) + 2B_{FV}(x)/\left(T - T_o(x) + \sqrt{[(T - T_o(x))^2 + C_{FV}(x)T]}\right), \quad (S2)$$

with $A_{FV}(x)$ again representing the high-temperature limit, and $B_{FV}(x)$, $C_{FV}(x)$ and $T_o(x)$ latent variables related to free volumes' properties like percolation, size and numbers.

Beyond the Free Volume and Adam-Gibbs models, many other theories have been proposed to describe the viscous flow of liquids. Among those, some are empirical like the Tamman-Vogel-Fulcher (TVF) equation, or semi-empirical like the Avramov and Milchev (AM) (32) model. We can also cite the MYEGA model (33) that directly derives from eq. S1. The TVF equation is

$$\log\eta(T, x) = A_{TVF}(x) + B_{TVF}/(T - T_1), \quad (S3)$$

with A_{TVF} , B_{TVF} and T_1 adjustable parameters. The AM model proposes the equation

$$\log\eta(T, x) = A_{AM}(x) + (12 - A_{AM}(x))\left(T_g(x)/T\right)^{m(x)/(12 - A_{AM}(x))}, \quad (S4)$$

with A_{AM} a pre-exponential terms proportional to $\log\eta(T \rightarrow \infty)$, $T_g(x)$ and $m(x)$ the melt glass transition temperature and fragility. The MYEGA equation can be written in a quite similar form:

$$\log\eta(T, x) = A_e(x) + (12 - A_e(x))\left(T_g(x)/T\right)e^{\left(m(x)/(12 - A_e(x)) - 1\right)\left(\frac{T_g(x)}{T} - 1\right)}, \quad (S5)$$

with A_e a pre-exponential term proportional to $\log\eta(T \rightarrow \infty)$ that was taken as equal to that in the Adam-Gibbs theory (eq. S1) because the MYEGA equation is a daughter product of the Adam-Gibbs theory. Equations S3 to S5 remain empirical or semi-empirical as unlike the Adam-Gibbs or Free Volume equations, they are not expressed in terms of measurable physical quantities, such as heat capacity. However, they still represent the viscosity dependence upon temperature well (Fig. S3). The MYEGA and AM equations also allow leveraging knowledge of melt T_g and fragility as predicted by the deep learning framework.

Deep learning framework implementation

The deep learning *framework* presented in Fig. 1A was implemented in Python using the PyTorch library. It takes four inputs: the mole fractions of SiO_2 , Al_2O_3 , Na_2O and K_2O . These are fed into to a neural network composed of n hidden layers, each one having k activation units (a.k.a neurons). Having explored various alternatives, we adopted the now-popular rectifier function (34), so that a neuron receiving input x returns output $y = \max(0,x)$. The outputs of this core network were fed to two final linear layers. The first output layer returns vectors that we regard as Raman spectra, calculated from the linear sum of the last neural network hidden layer. The second output linear layer returns 16 different values, which we interpret as:

- the parameters A_e , A_{AM} , A_{CG} and A_{TVF} (eqs. S1 to S5), as well as the coefficients B_1 to B_3 and C_1 to C_3 of the Sellmeir equation (see eq. S6) for the calculation of the glass refractive index n are directly given by the linear outputs; and
- the natural logarithms of $S^{conf}(T_g)$, C_{CG} , T_g , T_o , T_l , the melt fragility m , and the glass density d .

The use of the logarithm in the latter case was inspired by a strategy proposed by (29), and ensures that quantities are assigned positive values in accordance with their physical meaning.

We also find it aids rapid convergence during training. One trick for this method to work is to properly set the initial values of the biases of the last output layer to realistic values when creating the network.

Neural network predictions can be used in the physical theories embodied by equations S1 to S5 to obtain predictions of melt viscosity, and provide directly different values like glass density, glass transition temperature or Raman spectra. Furthermore, the coefficients B_1 to B_3 and C_1 to C_3 are used to predict the refractive index at given wavelength, $n(\lambda)$, via the Sellmeier equation:

$$n(\lambda) = \sqrt{1 + \frac{B_1 \times \lambda^2}{\lambda^2 - C_1} + \frac{B_2 \times \lambda^2}{\lambda^2 - C_2} + \frac{B_3 \times \lambda^2}{\lambda^2 - C_3}}. \quad (S6)$$

Our neural network therefore allows us to input compositions and obtain predictions for:

- melt viscosity, within five distinct theoretical frameworks,
- glass transition temperature,
- latent variables like configurational entropy and fragility,
- glass density,
- glass refractive index as a function of wavelength, and
- glass Raman spectra.

These predictions depend on a large number of tuneable parameters embedded within the neural network. During network training, we use our database of observed glass properties to optimize these parameters, seeking good average predictive performance.

The architecture of the hidden layers was optimized via a random search (35). This allowed us to observe how the number of layers and hidden units affect the generalization ability of the deep learning framework (Fig. S2). After training 2000 artificial neural networks under the same conditions on the same datasets, we observed that moderately deep networks with 3 to 5 layers and 200-300 units per layer perform best (Fig. S2); accordingly, best performance is in general reached with more than 1000 neurons. The dropout method, which consists in turning off p percent of neurons per layer at each training iteration in order to prevent overfitting (13), slightly helps preventing overfitting but is not a critical feature (Fig. S2). From this random search, we selected the 10 best neural networks with the lowest error on the validation data subset. All reported predictions are made using those 10 best neural networks, again their combination allowing better estimations to be made, following the well-known bagging principle (15).

Training

During training, we monitored the least-square deviations between measurements and predictions for the viscosity from eqs. S1 to S5 (see below) as well as density, optical and Raman spectra. We further added loss functions for known viscous T_g and $S^{conf}(T_g)$ values in the dataset $D_{viscosity}$. This enables better-constrained estimates of $S^{conf}(T_g)$. This parameter is usually hard to calculate as eq. S1 does not have a well-defined, unique solution, due to strong correlations between B_e and $S^{conf}(T_g)$ originating from the intervention of the intrinsic entropy S_c^* in both B_e and $S^{conf}(T_g)$ (7). The present approach was introduced to make the network less sensitive to this correlation. For the same reason, the network does not predict directly B_e ; this term is instead calculated as

$$B_e = (12.0 - A_e(x)) \left(T_g(x) S^{conf}(T_g, x) \right). \quad (S7)$$

A similar strategy was adopted to calculate B_{FV} and B_{TVF} from the other parameters.

Batch training was performed monitoring the global loss on the *training* and *validation* data subsets. Early stopping (14) was used to avoid overfitting: when the global loss function on the *validation* data subset stopped decreasing for more than 50 epochs, training was stopped and the network presenting the best validation loss was saved. This, combined with the other strategies (dropout, bagging and multitask learning), avoided overfitting as much as possible despite a limited and sparse dataset.

Statistical Analysis

Following the bagging method, we make predictions based on the mean of the ten best neural networks obtained from the random exploration of the network architecture (see above). This improves generalization, and also allows a statistical analysis of the influence of the network size on the predictive ability of the deep learning framework (see previous section as well as Fig. S2). The influence of the dataset size was explored through training ten neural networks with the same architecture with different *training* subset of $D_{viscosity}$. Results of this experiment thus represent the mean of those 10 different neural networks. Finally, the correlation between the different predicted parameters was explored using the Spearman correlation coefficient that allows observing non-linear correlations between different variables (see Supplementary Notebook).

Supplementary Text

The structure of aluminosilicate melts

Aluminosilicate melts are a disorganized network of SiO_2 and AlO_2 tetrahedral units, with bonds formed between them; some of these bonds get disrupted by network modifier elements like alkali and alkaline-earth elements. The later also compensate for the deficit of charge around AlO_2 tetrahedral units. In such melts, viscous flow involves cooperative movements of the tetrahedral units, facilitated by the presence of network modifiers cations like alkali and alkaline-earth elements. Rapid cooling of such melts allows crossing the glass transition, where the melt structure is frozen-in. Melt structure close to the glass transition can thus be observed in their glasses via, e.g., ^{29}Si Nuclear Magnetic Resonance or Raman spectroscopy.

It is important to consider melt structure and the role of elements because this directly determines melt properties. For instance, in the present case, variations in $S^{\text{conf}}(T_g)$ and m can be understood once we consider that aluminum and non-network former metal cations have important and complex roles in the melt (see review of 10). In Al-free silicate glasses, network modifier alkali cations break Si-O-Si bonds, forming alkali channels percolating in the disrupted SiO_2 tetrahedral network as described by the Modified Random Network (MRN) (36–40). Adding aluminum, entering as network forming AlO_4^- tetrahedral units in glasses, changes this picture: alkali metals switch their role from network modifiers to charge compensators of AlO_4^- tetrahedra to ensure charge balance (23). Alkali distribution still is non-random, but localized in compensator channels as described by the Compensated Continuous Random Network (CCRN) model (4, 41). In term of properties, changing the role of alkali metals from network modifiers to

charge compensators of Al results in an average decrease in $S^{conf}(T_g)$, particularly marked in the case of potassium compositions (Fig. 3a,b). MRNs thus may generally present higher $S^{conf}(T_g)$ than CCRNs, at least for alkali aluminosilicates. Furthermore, mixing alkalis results in different MAE effects as the alkalis reside in MRN (for Al-free or poor melts, Fig. 3c) or CCRN (in Al rich melts, Fig. 3f). In the former case, mixing Na and K induces an excess entropy of mixing caused by the hindering of the diffusions of alkali cations in modifier percolation channels (41). In the latter case, variations in $S^{conf}(T_g)$ with the Na/K ratio are close to a linear mechanical mixing of two sub-networks (Na-Al-Si-O and K-Al-Si-O subnetworks) because K and Na occupy different environments and do not really interact upon mixing (4, 24).

The difficulty of modeling aluminosilicate melts properties

The AG theory assumes that melt viscous flow occurs through cooperative re-arrangement of molecular subunits, and such events have been identified via high-temperature ^{29}Si NMR spectroscopy (42) and even direct observations (43, 44) in silicate melts. The structure of simple Al-free silicate melts has been related to $S^{conf}(T_g)$ (7) and C_p^{conf} (45), and this was leveraged to calculate melt viscosity in the ternary $\text{Na}_2\text{O}-\text{K}_2\text{O}-\text{SiO}_2$ system with an unrivaled precision of 0.2 log Pa·s (7). However, it is very difficult to extend to more complex composition like aluminosilicates because of the many new degrees of complexity generated by the addition of one critical elements like Al.

Indeed, it is actually difficult to experimentally validate models of melt structure for aluminosilicate compositions, representative of most natural and industrial glasses. In such compositions, ^{29}Si NMR spectroscopy, which usually brings the necessary information to

quantify the connectivity of SiO_2 tetrahedral units, becomes blind due to Si-Al interactions causing significant signal broadening; signal interpretation relies on various hypothesis and back-end models of melt structure (46). Raman spectroscopy, another method to explore glass structure, does not solve this problem as it cannot be calibrated against reliable NMR data to distinguish the molecular subunits in the glasses. Furthermore, the aluminum content of the melt also affects interactions between, and the environment of the metal cations, as well as Al-Si ordering, Al coordinance and the potential presence of three-fold coordinated oxygen (see 10 for a review). Such problems severely affect our ability to construct models in the presence of aluminum, and strongly question the theoretical viability of proposed models based on untested structural calculations (e.g. 8). This complexity pushed existing models to simply link chemical composition of aluminosilicate melts to their viscosity using a set of polynomial equations (5, 47, 48). A more complex model was proposed by Starodub et al. (8) for the system $\text{Na}_2\text{O}-\text{K}_2\text{O}-\text{Al}_2\text{O}_3-\text{SiO}_2$. They proposed an associate-solution model for calculating the structure of melts in this system. Their model is interesting but one should keep in mind that the structural calculations was not validated by experimental data for aluminosilicate melts, and therefore there may be biases due to our inability to determine the structure of aluminosilicate compositions. Other biases also come from the absence of viscosity data selection in their model calibration. This can severely damage the model results since melt viscosity can be difficult to measure with precision, particularly at undercooled conditions close to the glass transition temperature.

Internal consistency of the deep learning framework

The ability to predict the melt fragility m allows further testing of the internal consistency of our deep learning framework predictions. Indeed, experimental data indicate that a direct correlation between m and the ratio between the configurational heat capacity at T_g and $S^{conf}(T_g)$, $C_p^{conf}(T_g)/S^{conf}(T_g)$ shall be observed (49, 50). The deep learning framework predicts this linear correlation (Figure S5), albeit some scattering that most probably arises from the way melt and glass C_p^{conf} values are determined in the model. Indeed, the deep learning framework calculates $C_p^{conf}(T_g)$ as $C_p^{liquid}(T_g) - C_p^{glass}(T_g)$, with $C_p^{liquid}(T_g)$ predicted from the model of (51) and $C_p^{glass}(T_g)$ calculated from the Dulong-Petit limit of $3R$, with R the perfect gas constant. Models of $C_p^{liquid}(T_g)$ and the theoretical calculation of $C_p^{glass}(T_g)$ do not yield exact values, and this most probably explains the moderately good correlation between predicted m and $C_p^{conf}(T_g)/S^{conf}(T_g)$ values in the present work in comparison of other studies that used experimental values of C_p . In turn, this result indicates a critical need for better heat capacity models of aluminosilicate melts.

Deep learning for small dataset?

Using the viscosity dataset, we tested the effects of dataset size and network architecture on the deep learning framework predictive abilities. Experimental data are often scarce because they are difficult to obtain, with each experiment requiring significant work, time, and potentially funds. The common idea is that this usually prohibits the use of “data-thirsty” machine learning methods. The present datasets are small, even extremely small (e.g. that of Raman spectra), raising the question: are we not simply overfitting or even encoding our data? We tested this by training the neural networks on different training data subsets with variable size and monitoring the RMSE loss on the same unseen test data subset (Fig. S2A). As expected, the deep learning framework predictive ability directly depends on the dataset size. Results become interesting

only after reaching a threshold of around ~ 80 different compositions within the training data subset. This yields testing RMSE lower than $0.6 \log \text{Pa}\cdot\text{s}$ on the *testing* $D_{viscosity}$ subset (Fig. S2A), a reasonable achievement as existing parametric models have RMSE values higher than $0.6 \log \text{Pa}\cdot\text{s}$ on the full 10^0 - $10^{15} \text{ Pa}\cdot\text{s}$ range (5, 47, 48). Lower numbers of compositions in the *training* subset result in the deep learning framework constantly over-fitting the data. With more data, the lowest achievable RMSE probably lies around $0.36 \text{ Pa}\cdot\text{s}$ as shown by the evolution of the training RMSE (Fig. S2A) and reflects errors affecting the dataset (see below).

Aside from the number of training data, the network architecture has a direct effect on the deep learning framework predictive ability. For small datasets, common but outdated advice to avoid overfitting is to use networks with a limited number of activation units and layers. In our case, this is not true. More than a thousand hidden activation units are necessary to achieve good validation and testing RMSE on the viscosity dataset (Fig. S2). More precisely, network with more than 3 hidden layers provide better predictive abilities (Figure S2), confirming that deep networks perform better than shallow ones even on small datasets.

Can we detect a lowest achievable limit for the deep learning framework predictive error?

Interestingly, the training RMSE on viscosity is always of $\sim 0.36 \log \text{Pa}\cdot\text{s}$ (Fig. S2A), regardless of the training data subset size. Artificial neural networks are extremely flexible and very prone to overfit the training data subset. Therefore, the fact that neural networks trained on a small number of compositions clearly overfit the training data subset but do not provide viscosity predictions better than $\sim 0.36 \log \text{Pa}\cdot\text{s}$ brings an important information: this places a lowest achievable RMSE at $\sim 0.36 \log \text{Pa}\cdot\text{s}$ on the present dataset. We infer that this limit reflects

contributions from experimental errors affecting viscosity measurements and chemical composition values, those errors varying between different laboratories, as well as from the accuracy of the glass and melt heat capacity determination (see above).

Configurational entropy at the glass transition: chemical and topological contributions

Deep learning framework results show large, non-linear variations of $S^{conf}(T_g)$ with glass composition. Such variations are better understood when considering that $S^{conf}(T_g)$ should be considered as originating from various sources, of topological and chemical nature. As a result, it varies in a complex way with melt chemistry, and those variations are difficult to rationalize without making severe simplifications and/or assumptions. Even for the simple SiO₂ glass, not all the configurational entropy can be assigned to a topological origin with certainty. Indeed, the Raman signal of the A1 vibrational doublet at high frequency arising from symmetric stretching of the SiO₂ tetrahedral units is split in two components, which fractions represent 0.68 and 0.32 of the A1 signal (20); assuming an ideal mixing of the two vibrators, one finds a Raman-derived $S^{conf}(T_g) = 5.2 \pm 0.4 \text{ J mol}^{-1} \text{ K}^{-1}$. This value is surprisingly realistic, very close to the $5.1 \pm 2 \text{ J mol}^{-1} \text{ K}^{-1}$ value returned from calorimetric measurements (52) and not very different from the $8.3 \pm 2.8 \text{ J mol}^{-1} \text{ K}^{-1}$ value calculated from the SiO₂ viscosity data (7). The same treatment for the nepheline glass yields a value of $4.6 \pm 0.7 \text{ J mol}^{-1} \text{ K}^{-1}$, close to the calorimetric value of $4.85 \text{ J mol}^{-1} \text{ K}^{-1}$, this being explained by the high ordering of the Si-Al distribution in nepheline (53) and hence a limited chemical mixing of various units. However, as soon as chemical mixing becomes important, such calculations fail. This is shown by comparing the value of such calculation for albite, which yields a Raman-derived $S^{conf}(T_g) = 3.8 \pm 1.1 \text{ J mol}^{-1} \text{ K}^{-1}$, a value much lower than the calorimetric one equal to $9.2 \pm 2 \text{ J mol}^{-1} \text{ K}^{-1}$ (52). Assuming that Na

repartition in CCRN is not accompanied by any mixing effect (this seems reasonable as Na is the only cation that can occupy such environment), this difference can be assigned to Al-Si mixing. Assuming ideal mixing, one can calculate it and add it to the $3.8 \pm 1.1 \text{ J mol}^{-1} \text{ K}^{-1}$ value to finally obtain a value of $8.46 \pm 1.1 \text{ J mol}^{-1} \text{ K}^{-1}$, a value in much better agreement with the measured one. Such analysis may not be easily extrapolated to other compositions, but rather is an interesting exposition of the source of entropy affecting the melt configurational entropy at the glass transition, and, hence, the variations of its viscosity. In turn, the complexity of $S^{\text{conf}}(T_g)$ indicates that values for different glasses do not necessarily embed interesting information about their relative ordering. Indeed, while two glasses may present similar $S^{\text{conf}}(T_g)$, their decomposition may yield very different topological and mixing contributions, making any attempt in discussing glass structure using $S^{\text{conf}}(T_g)$ pointless.

Links between the Adam-Gibbs and the Free Volume theories?

The trans-theoretical character of the present deep learning framework allows adopting a new vision about problems such as 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 intrinsic entropy S_c^* . 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 eqs. S1 and S2. Indeed, B_{CG} embeds some structural information because it depends on v^* :

$$B_{FV} = v^* z_o, \quad (S8)$$

where z_o is an adjustable parameter. Similarly, the ratio $B_e/S^{conf}(T_g)$ embeds molecular subunit lengthscale information as

$$B_e/S^{conf}(T_g) = [\Delta\mu z^*(T_g)] / R, \quad (S9)$$

with $\Delta\mu$ the energy barriers opposed to the rearrangement of molecular subunits of size $z^*(T_g)$, and R the perfect gas constant. We can therefore 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 B_{FV} and $B_e/S^{conf}(T_g)$ correlate very well with R_{Raman} ($r_s = 0.987$ and 0.985 respectively), which provides information about the network connectivity; the higher R_{Raman} is, the higher the interconnection between polyhedral units, thus the higher the 3D network topology(54). 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 (55, 56). More generally, the links between B_{FV} , $B_e/S^{conf}(T_g)$ and R_{Raman} support the general hypothesis that melt viscous flow occurs when a critical molecular lengthscale is reached. This lengthscale can be determined from Raman spectra (Fig. 2C,D) and strongly influences the glass transition temperature T_g (Fig. 2A). In detail, entropic effects (like the excess of entropy resulting from the MAE) also affect T_g (7 and references cited therein)but their influence remains limited compared to that of the polyhedral network topology (Fig. 2A). On the other hand, 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 $m \propto C_p^{conf}/S^{conf}(T_g)$ (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 T_g . Following from this idea, it is possible to relate m to ongoing density fluctuations inherited from dynamic heterogeneities formed at supercooled T, i.e. $T_{\text{liquidus}} < T < T_g$ (57). 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 AX_2 glasses like SiO_2 (58), the Modified Random Network for silicate liquids (37) and the Continuous Compensated Random Network (4, 41) 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 and hence viscosity.

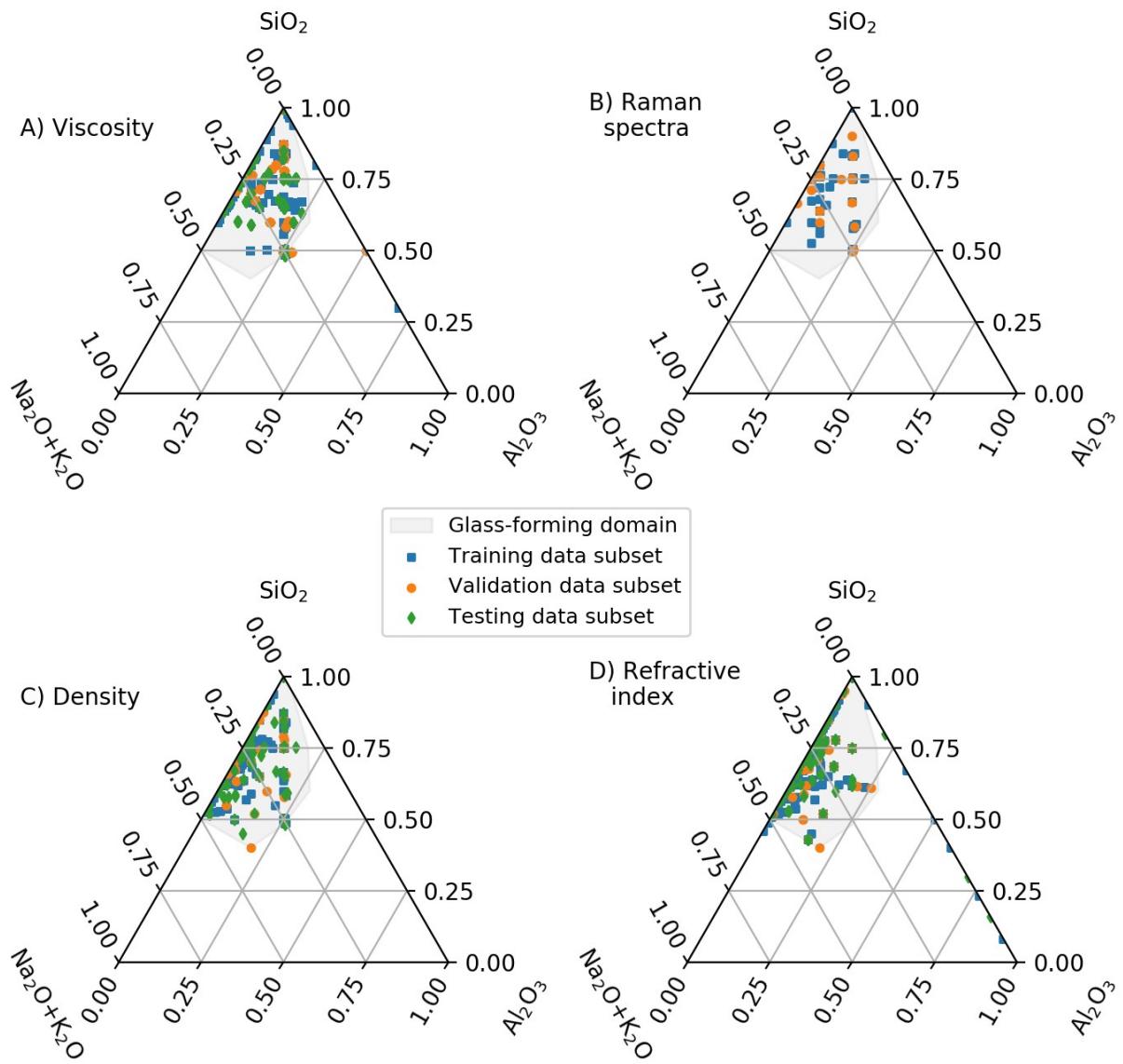


Fig. S1. Viscosity (A), Raman (B), density (C) and refractive index (D) datasets used in this publication. Each symbol corresponds to a composition. The glass-forming domain at usual laboratory cooling rates is indicated in grey.

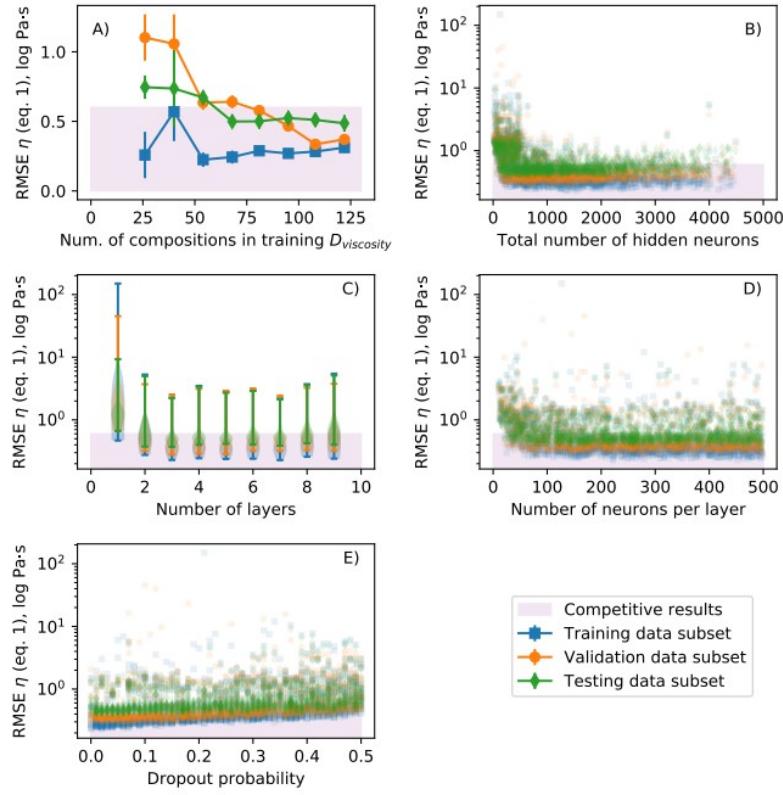


Fig. S2. Network architecture and dropout influence on deep learning framework

performance. RMSE between viscosity predictions made using the Adam-Gibbs model (eq. 1) and measurements in training, validation and testing data subsets as a function of the number of compositions in the training data subset (a), of the total number of hidden activation units (a.k.a. neurons, b), of the number of hidden layers (c), of the number of neurons per layer (d) and of the dropout probability (e). 2,000 neural networks were randomly selected and trained to obtain those results. Subplot (c) is a violin plot with extreme values showed. Subplots (b), (d) and (e) are scatter plots in which each slightly transparent symbol corresponds to a given neural network; less transparency is directly indicative of a higher number of models for a given X-Y value. Deep network still generalizes better than shallow one on this small problem. Overfitting is very limited limited but still present, as shown by systematically slightly lower errors on the training data subset.

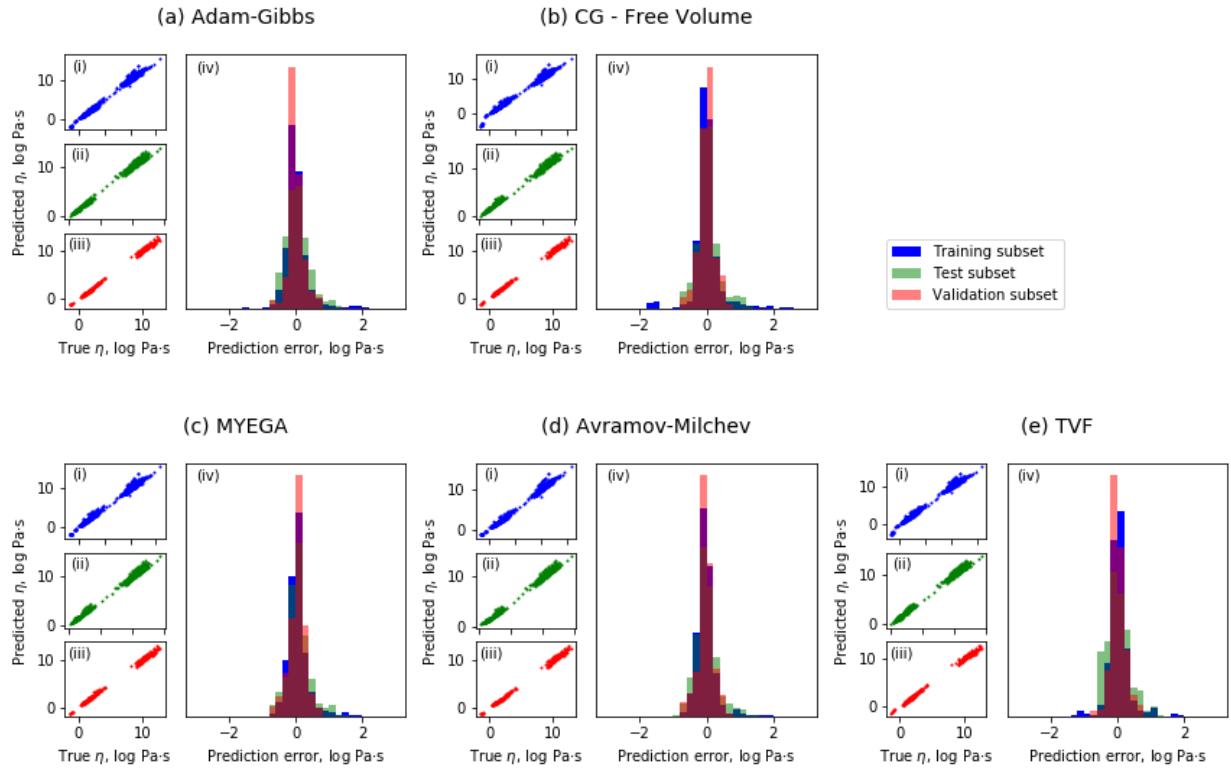


Fig. S3. Comparison between predicted and measured viscosity in the $\text{Na}_2\text{O}-\text{K}_2\text{O}-\text{Al}_2\text{O}_3-\text{SiO}_2$ system. Predictions can be made using theories like Adam-Gibbs (a) and free volume (b), or empirical equations like MYEGA (c), Avramov-Milchev (d), and Tamman-Vogel-Fulcher (e). See table S3 for RMSE and text for the equations.

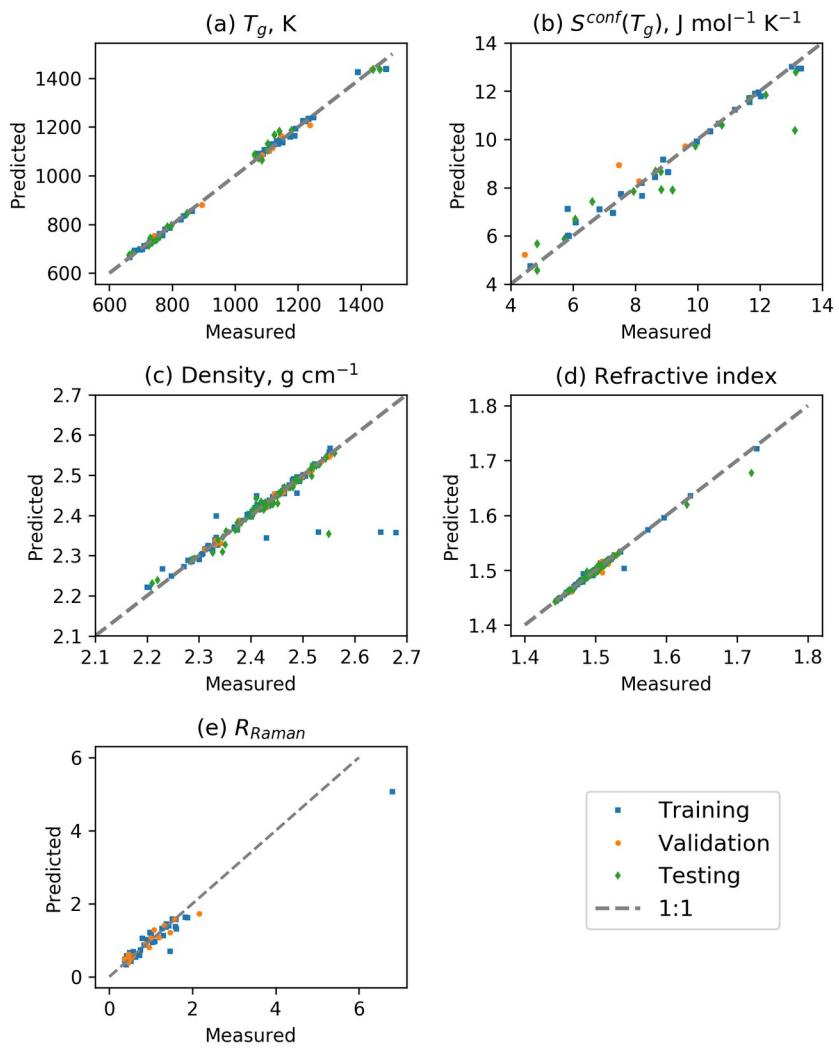


Fig. S4. Comparison between predictions and observed viscous glass transition temperature (a), $S^{conf}(T_g)$ (values from 4, 7, 19, 20, 23, 24) (b), density (c), refractive index (d), and R_{Raman} (e). See table S3 for RMSE.

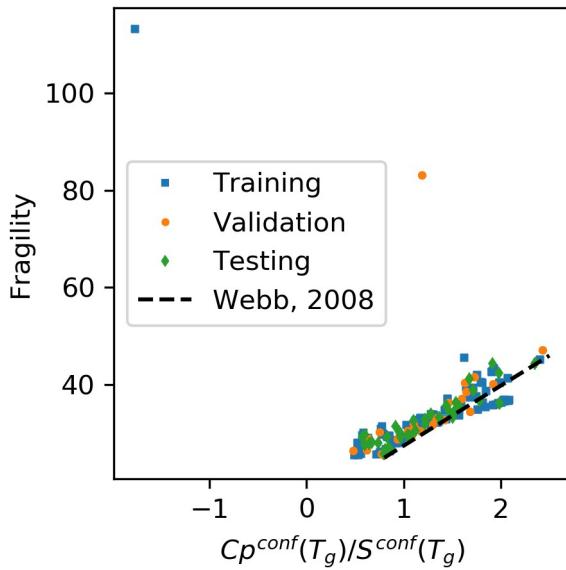


Fig. S5 Glass fragility versus melt $C_p^{\text{conf}}(T_g)/S^{\text{conf}}(T_g)$ ratio. Symbols are predictions of the deep learning framework on the different subsets of the $D_{\text{viscosity}}$ dataset. The back dotted line is the relationship observed by (49) using experimental heat capacity data. Except two extreme outliers that corresponds to Al_2O_3 - SiO_2 melts with more than 30 mol% Al_2O_3 , a general good agreement is observed. This validates the internal consistency of the deep learning framework. The scatter indicates that a better agreement could be obtained with using a better heat capacity model for aluminosilicate melts. Such C_p calculation could be integrated in the present deep learning framework in the future albeit the acquisition of additional C_p data for ternary and quaternary aluminosilicate melts.

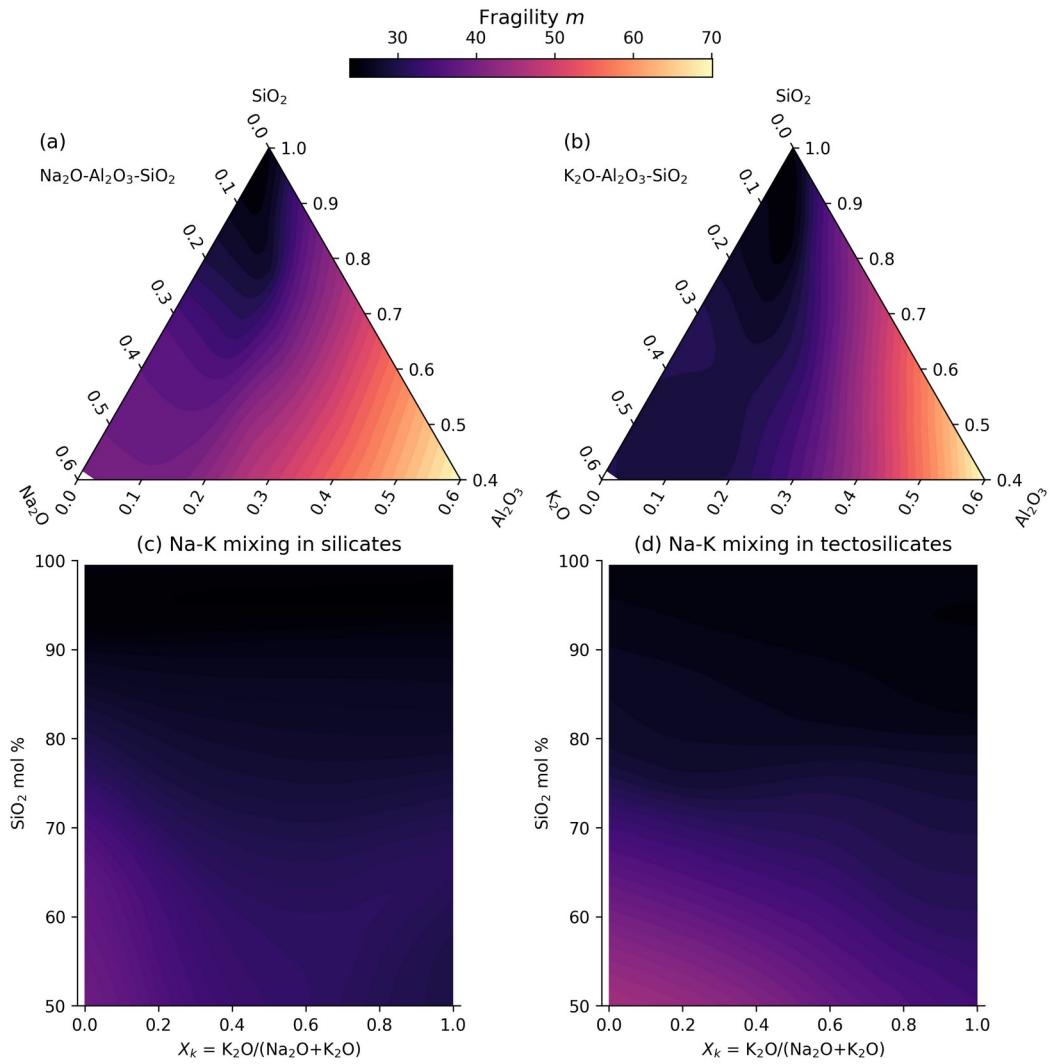


Figure S6: Melt fragility variations with composition. Fragility is represented in the glass forming domains of the ternary sodium (a) and potassium (b) aluminosilicate systems, as well as as a function of the silica fraction and the potassium to total alkali ratio of silicate (c) and tectosilicate (d) melts. No MAE is observed on melt fragility, which depends largely on melt silica content.

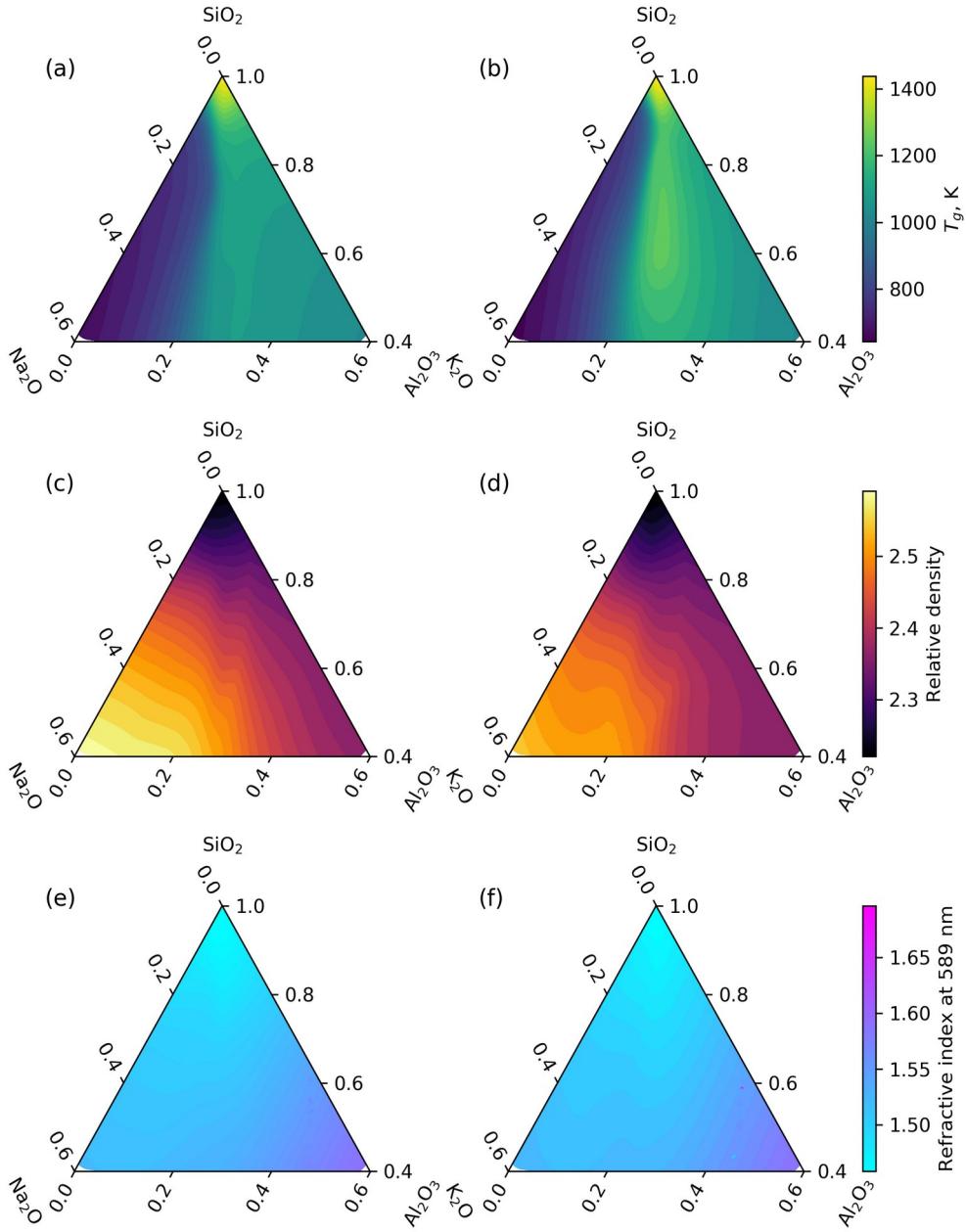


Fig. S7 Deep learning framework predicted variations in glass transition temperature, relative density and refractive index in the ternary Na and K aluminosilicate diagrams.

Glass name		%SiO₂	%Al₂O₃	%K₂O	%Na₂O	Density, g cm⁻¹
KA80.05	nom. mol%	80.00	5.00	15.00	0.00	
	nom. wt%	71.40	7.60	21.00	0.00	
	an. wt%	74.85(44)	7.56(14)	15.12(22)	0.00(4)	2.320(1)
KA72.07	nom. mol%	72.00	7.00	21.00	0.00	
	nom. wt%	61.60	10.20	28.20	0.00	
	an. wt%	61.41(28)	10.17(24)	27.42(33)	0.00(2)	2.408(1)
KA65.09	nom. mol%	65.00	8.75	26.25	0.00	
	nom. wt%	53.70	12.30	34.00	0.00	
	an. wt%	53.32(55)	12.49(43)	31.68(28)	0.00(3)	2.451(10)
NA65.09	nom. mol%	65.00	8.75	0.00	26.25	
	nom. wt%	60.79	13.89	0.00	25.32	
	an. wt%	61.75(45)	13.67(24)	0.03(2)	24.56(73)	2.472(4)
NA58.10	nom. mol%	58.00	10.50	0.00	31.50	
	nom. wt%	53.55	16.45	0.00	30.00	
	an. wt%	54.61(28)	16.42(18)	0.05(2)	28.92(37)	2.502(5)

Table S1. Composition of the synthesized glasses. Nominal (nom.) and analyzed (an.) compositions are reported. Standard deviations on measured values on 10 different spots (for EPMA measurements) or glass chips (for density measurements) are given in parenthesis (1 sigma confidence interval).

T, K	KA80.05	T, K	KA72.07	T, K	KA65.09	T, K	NA65.09	T, K	NA58.10
1013.1	9.10	921.5	9.37	941.3	9.55	834.0	9.01	827.3	10.10
1001.8	9.32	891.0	10.17	935.1	9.71	829.0	9.18	836.9	9.73
989.6	9.51	872.0	10.75	919.4	10.13	813.8	9.61	817.6	10.51
981.6	9.78	852.0	11.40	913.6	10.32	803.2	9.94	796.2	11.42
967.6	10.05			898.7	10.78	798.6	10.09	805.7	10.97
949.7	10.50			892.0	10.96	787.5	10.50	847.1	9.36
940.3	10.83			882.1	11.28	779.9	10.81	855.9	9.09
928.2	11.05			867.3	11.86	773.7	11.07	828.3	10.01
918.4	11.32			855.4	12.29	772.9	11.07	834.4	9.81
905.1	11.63					762.6	11.52	787.4	11.80
896.6	11.92					756.3	11.74	777.4	12.36
						752.1	11.95		

Table S2. Viscosity measurements made on the synthesized glasses. Viscosity is in \log_{10} Pa·s and was measured using a creep apparatus following the protocol described in (4). Errors on temperature are lower than 0.3 K, and errors on viscosity lower, or equal to 0.03 \log_{10} Pa·s.

Data subset:	Training	Validation	Testing
Adam-Gibbs (eq. S3, log Pa·s)	0.33	0.22	0.35
Free Volume (eq. S4, log Pa·s)	0.36	0.23	0.36
TVF (eq. S5, log Pa·s)	0.36	0.25	0.35
MYEGA (eq. S7, log Pa·s)	0.36	0.25	0.34
Avramov-Milchev (eq. S6, log Pa·s)	0.34	0.24	0.34
Density (g cm⁻¹)	0.07	0.08	0.08
Raman spectra (%, LAD)	18	22	-
Refractive index	0.003	0.003	0.005

Table S3: Root-mean-square errors of the deep learning framework. RMSE calculated between measured and predicted melt viscosity, density, refractive index, except for Raman spectra where a different metric is used (median least absolute deviation LAD).