Geochemistry, Geophysics, Geosystems

Supporting Information for

Plagioclase-saturated melt hygrothermobarometry and plagioclase-melt equilibria using machine learning

Kyra S. Cutler[1], Mike Cassidy[1,2] and Jon D. Blundy[1]

[1] Department of Earth Sciences, University of Oxford, South Parks Road, Oxford, OX1 3AN, UK

[2] School of Geography, Earth and Environmental Sciences, University of Birmingham, Edgbaston, Birmingham, B15 2TT, UK

Contents of this file

Additional Supporting Information (see separate .xlsx spreadsheets)

Table S1. Calibration dataset for the thermometers, hygrometers, barometers, and the anorthite content model.

Table S2. Calibration dataset for the plagioclase-saturated classifier.

Table S3. Monte Carlo analytical uncertainty simulation input (ten experimental liquid compositions + weighted mean errors of electron microprobe glass oxide analyses) and output.

Introduction

This file contains supplementary text and figures omitted from the main manuscript. The supplementary text explains the error propagation used in the Mount St Helens application. The supplementary figures display the schematics of a random forest algorithm generating a prediction and 10-fold cross-validation, as well as significant geochemical relationships aiding algorithm prediction.

Text S1.

The ML H_2O-independent thermometer and T-dependent hygrometer provide an uncertainty value (standard deviation; SD) from the T or H_2O prediction of each individual glass compositional analysis. For n=50, a uniform distribution within the SD on every temperature or water content estimate predicted by the H_2O-independent thermometer and T-dependent hygrometer is sampled. For example, if the thermometer returns a value of 900 °C and an SD of ± 50 °C, 50 points are sampled between 850–950 °C according to a uniform distribution. All 50 temperature/water content estimates are then input into the T-dependent hygrometer or H_2O-dependent barometer for each glass analysis. The maximum absolute difference to the mean value is the maximum uncertainty associated with a given pre-eruptive water content/pressure estimate.

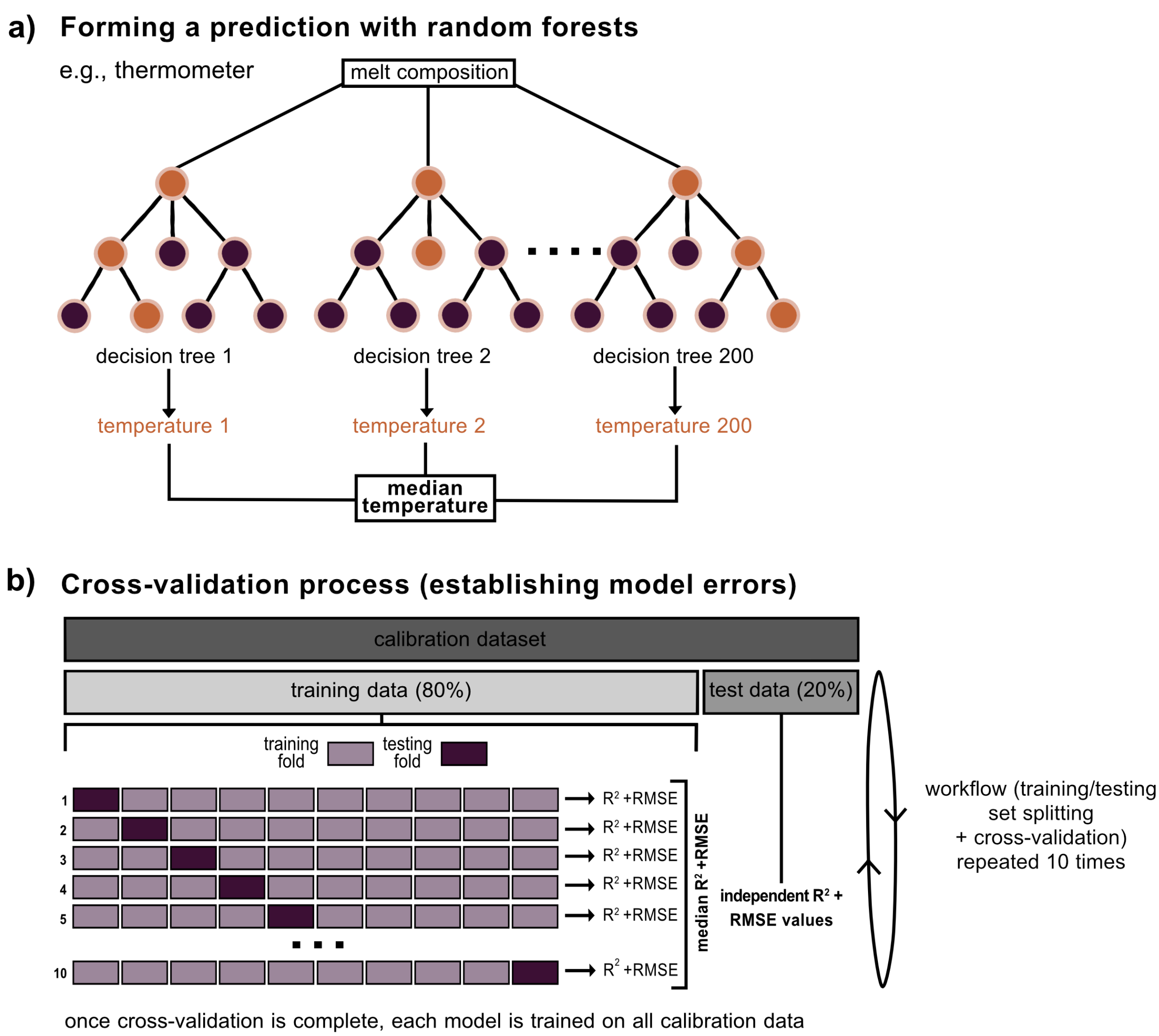

Figure S1. Summary diagrams of the random forest algorithm and the cross-validation process. **a**) Simplified diagram of a random forest thermometer displaying how a temperature prediction forms via averaging of multiple decision trees. **b**) Schematic of 10-fold cross-validation.

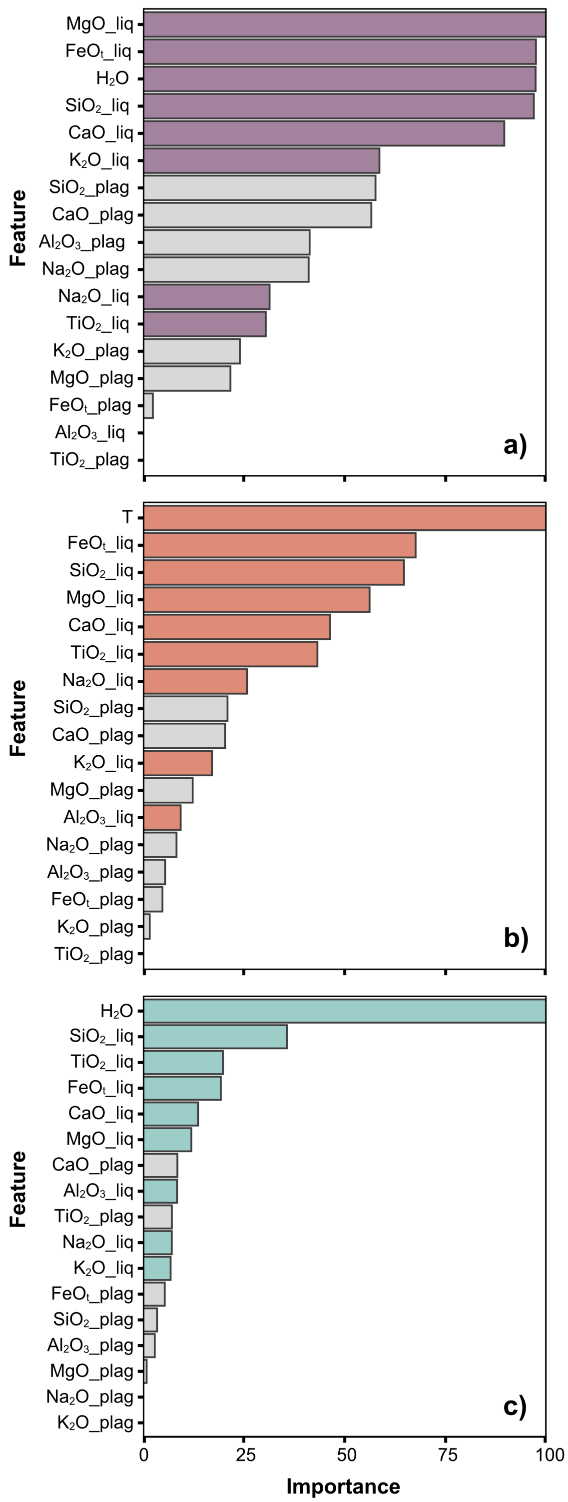

Figure S2. Plots of variable importance vs. input parameter for the H_2O-dependent thermometer (**a**), T-dependent hygrometer (**b**), and H_2O-dependent barometer (**c**), highlighting the most important variables used by the extratrees algorithm to make predictions. Grey bars represent plagioclase compositional inputs and coloured bars represent liquid compositional inputs along with additional parameters such as T or H_2O.

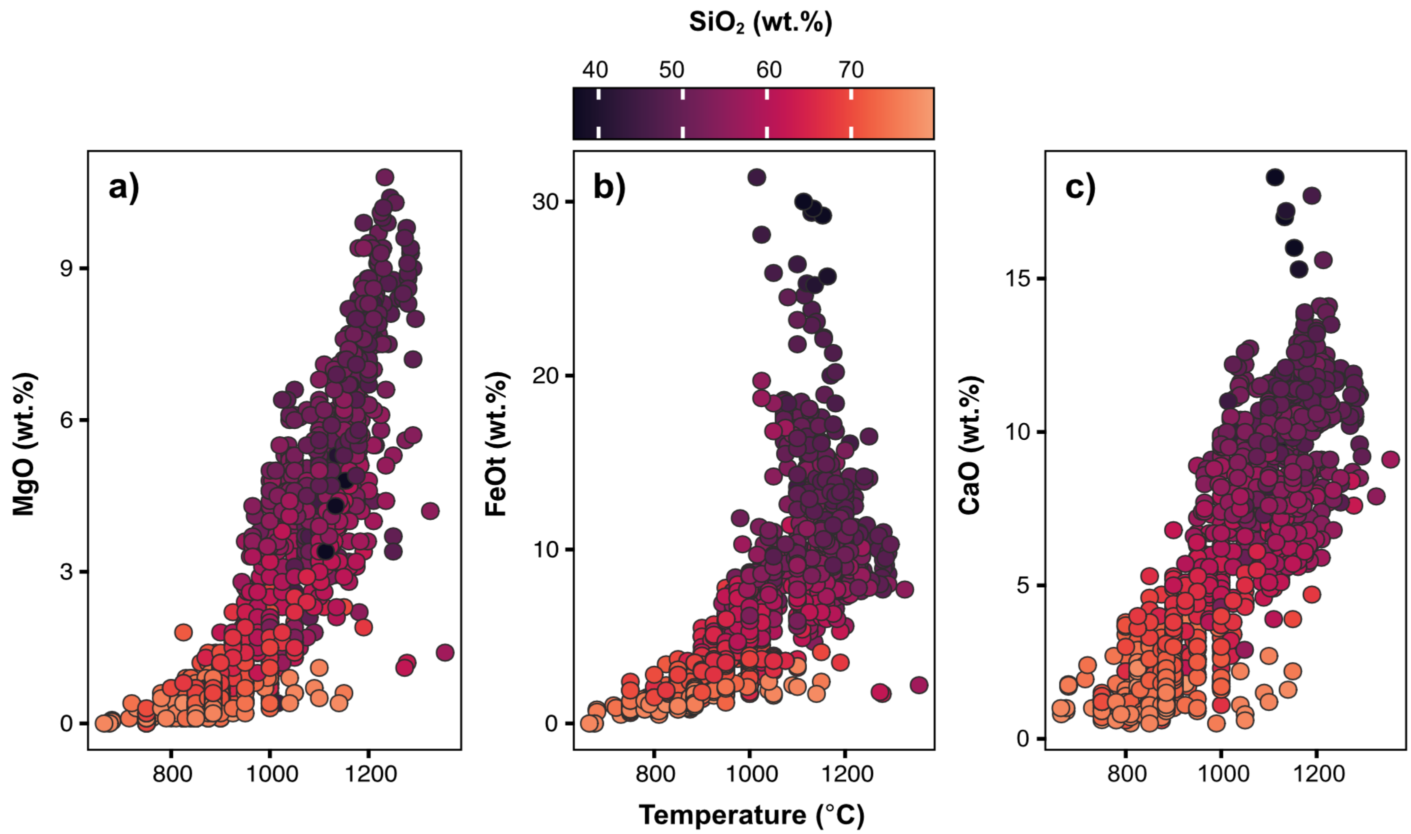

Figure S3. Plots illustrating the strong non-linearity between temperature and (**a**) MgO (wt.%), (**b**) FeOt (wt.%), and (**c**) CaO (wt.%) in the liquid. Colour coding reflects the SiO_2 (wt.%) of the liquid.

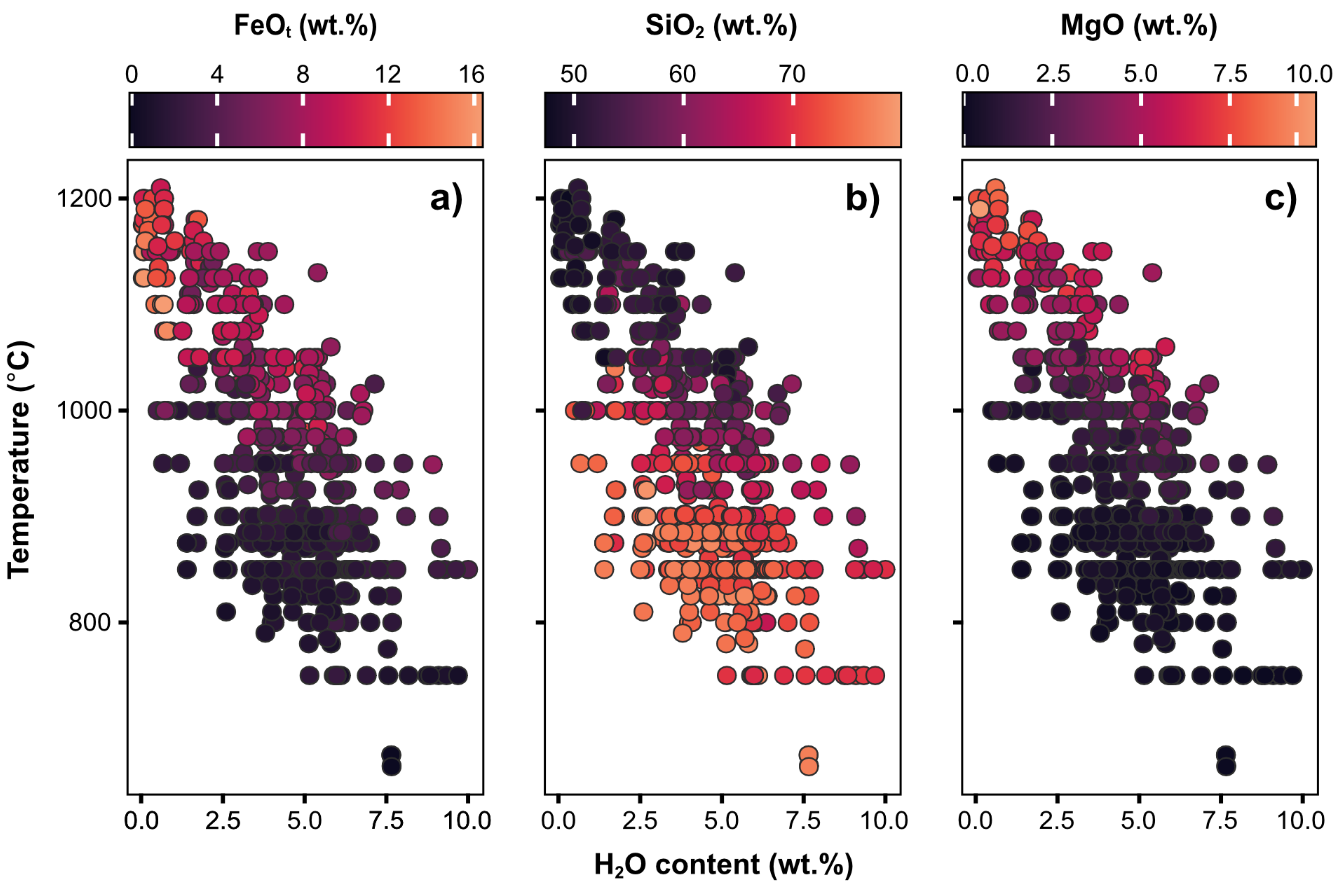

Figure S4. Temperature vs. H_2O (wt.%) of experimental glass compositions with colour coding showing the (**a**) FeO_t (wt.%), (**b**) SiO_2 (wt.%), and (**c**) MgO (wt.%) contents in the liquid.

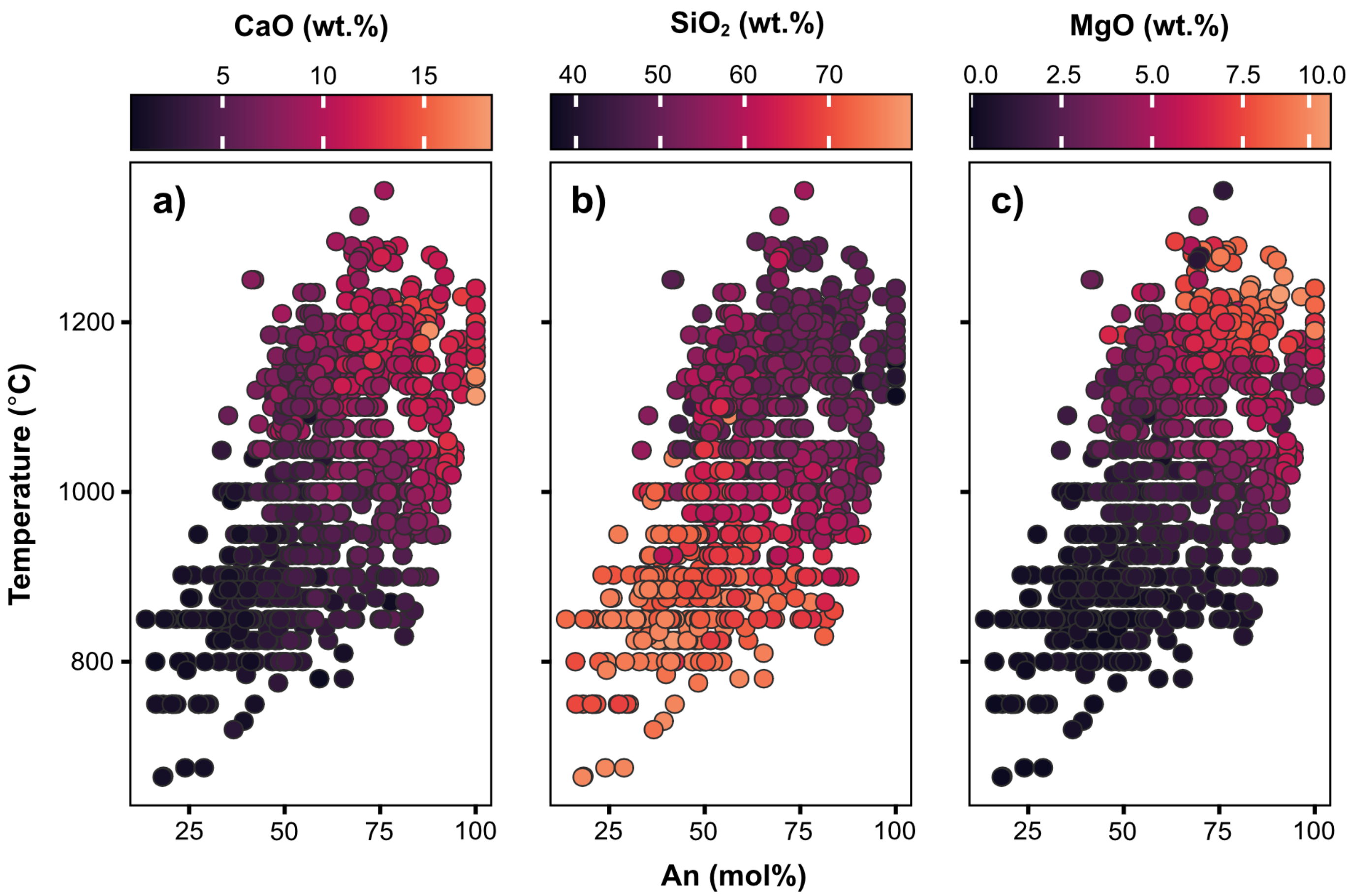

Figure S5. Temperature of experimental glass compositions vs. An content (mol%) of experimental plagioclase with colour coding showing the (**a**) CaO (wt.%), (**b**) SiO_2 (wt.%), and (**c**) MgO (wt.%) contents in the liquid.