POSTER 19

Machine Learning Workflows to Predict Crystallisability, Glass Forming Ability, mechanical properties of Small Organic Compounds.

Vijay K Srirambhatla,*[1] Blair Johnston[1,2], Alastair Florence[1,2]

[1]EPSRC ARTICULAR, University of Strathclyde, Technology Innovation Centre, 99 George Street, Glasgow, G1 1RD, UK
[2]EPSRC CMAC Future Manufacturing Research Hub, University of Strathclyde, Glasgow, G1 1RD, U.K.
E-mail: Vijay.Srirambhatla@strath.ac.uk

ARTICULAR
MEDICINES MADE SMARTER

Crystallisability

- **Crystallisability: is the ability of a compound to nucleate fast/medium/slow.**
- **Aim: To develop machine learning workflows to predict crystallisability.**
- In order to predict the crystallization propensity, experimental screening and image data collection was performed using the workflow shown in figure 1.

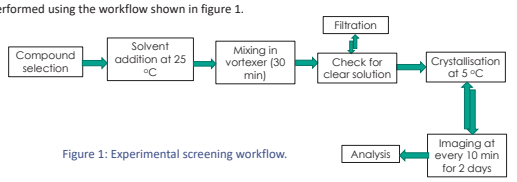


Figure 1: Experimental screening workflow.

The images from screening were grouped into 6 different categories.

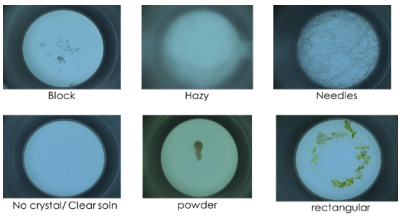


Figure 2: Classification of images from experimental screening.

- To develop a predictive model for crystallisability, image analysis was performed on the time series images by following the sequence of steps as shown in figure 6.

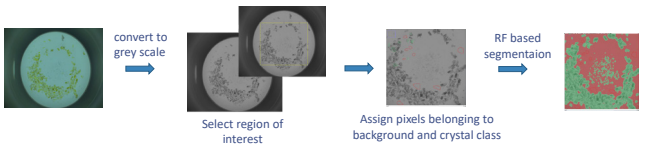
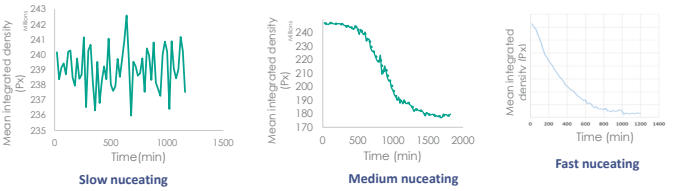


Figure 5: Sequential steps to classify time series images

- The RF model developed for segmenting one image was used to perform automatic segmentation of time series images.
- A plot of the decrease in mean background integrated intensity was used to classify datasets as fast, medium and slow nucleating.



- The results indicate a classification accuracy of ~73%.

Evaluation Results

Model	AUC	CA	F1	Precision	Recall
SVM	0.809	0.600	0.593	0.588	0.600
Random Forest	0.873	0.739	0.741	0.747	0.739
Neural Network	0.881	0.739	0.740	0.742	0.739

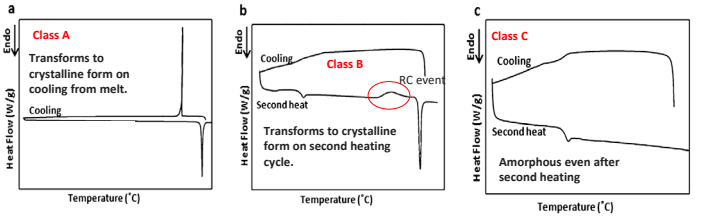
Predicted

Actual	Fast	Medium	Slow	Σ
Fast	25	6	4	35
Medium	6	30	4	40
Slow	7	3	30	40
Σ	38	39	38	115

Figure 9: Confusion matrix of the data to predict crystallisability.

Glass forming ability (GFA)

- Glass forming ability (GFA)1: The GFA of a material is the ease of vitrification of liquid on cooling.



Based on the experimental DSC curves the compounds are divided into three classes.

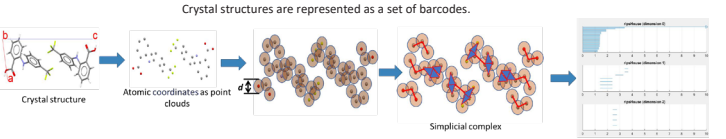
Mechanical Properties

Aim: to develop ML based method to predict mechanical properties of small organic compounds.

What information do we need ?

1. A way to represent crystal structures for machine learning.
2. Mechanical properties dataset (calculated/ measured).

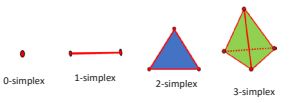
Crystal structures are represented as a set of barcodes.



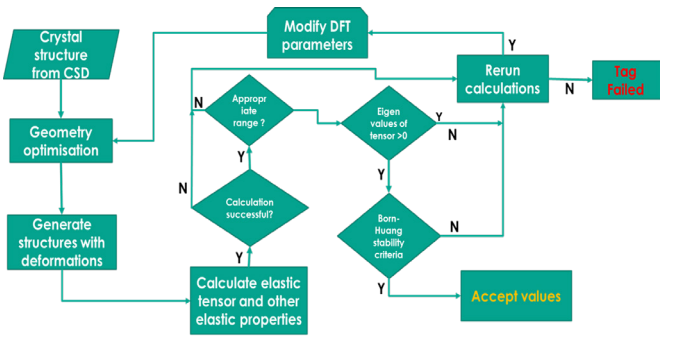
- The persistent barcodes **capture all molecular and intermolecular distances** in a crystal structure.

The simplicial complexes are generated using some simple rules where

- Each data point is represented as vertex **(0-simplex)**
- Every time two balls these balls intersect, you join an edge **(1-simplex)** between them.
- When three balls intersect then you draw a triangle **(2-simplex)**.
- When 4 balls intersect, we draw a tetrahedron **(3-simplex)**.

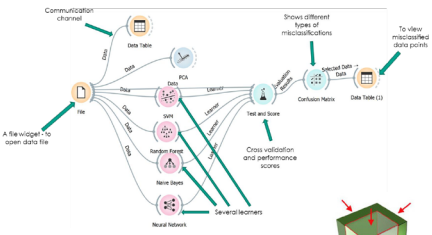


- Mechanical property calculations are using materials studio(CASTEP).
  Geometry optimised, Constant strain method.
  Functional: Generalized Gradient Approximation (GGA); Perdew, Becke and Ernzerhof (PBE)



Workflow to calculate mechanical properties computationally.

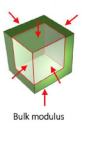
- The data from persistence homology + calculated mechanical properties were fed into ML workflow.

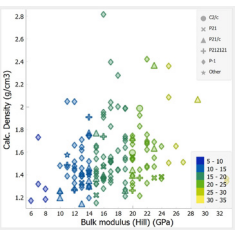


Generic ML workflow

Using **Bulk modulus** as target

Model	MSE	RMSE	MAE	R2
Evaluation Results				
Random Forest	0.106	0.325	0.195	0.995
Neural Network	0.628	0.707	0.687	0.954
Gradient Boosting	0.399	0.632	0.547	0.983





- **Bulk modulus** describes the response of the material to uniform pressure.
- The larger errors indicate we need more data points.
- The model was prepared using ~ **450** structures and their mechanical properties.

Summary

- ML workflows are developed to predict crystallisability, glass forming ability , and mechanical properties of small organic compounds.

Engineering and Physical Sciences Research Council · The Glasgow School of Art · University of Strathclyde Glasgow · Loughborough University · CCDC advancing structural science · BOOTH WELSH Integrated Engineering Services · SIEMENS · PERCEPTIVE ENGINEERING an Applied Materials company