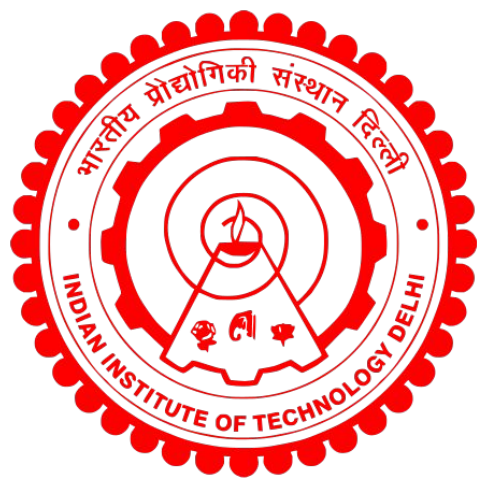
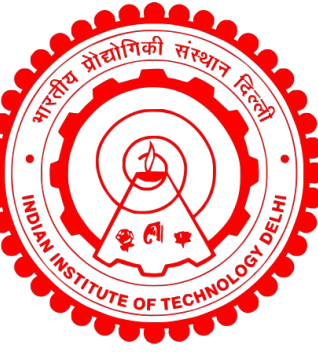


AI in Glass Strengthening

N. M. Anoop Krishnan

**Department of Civil and Environmental
Engineering
Yardi School of Artificial Intelligence
(Joint Appt.)**

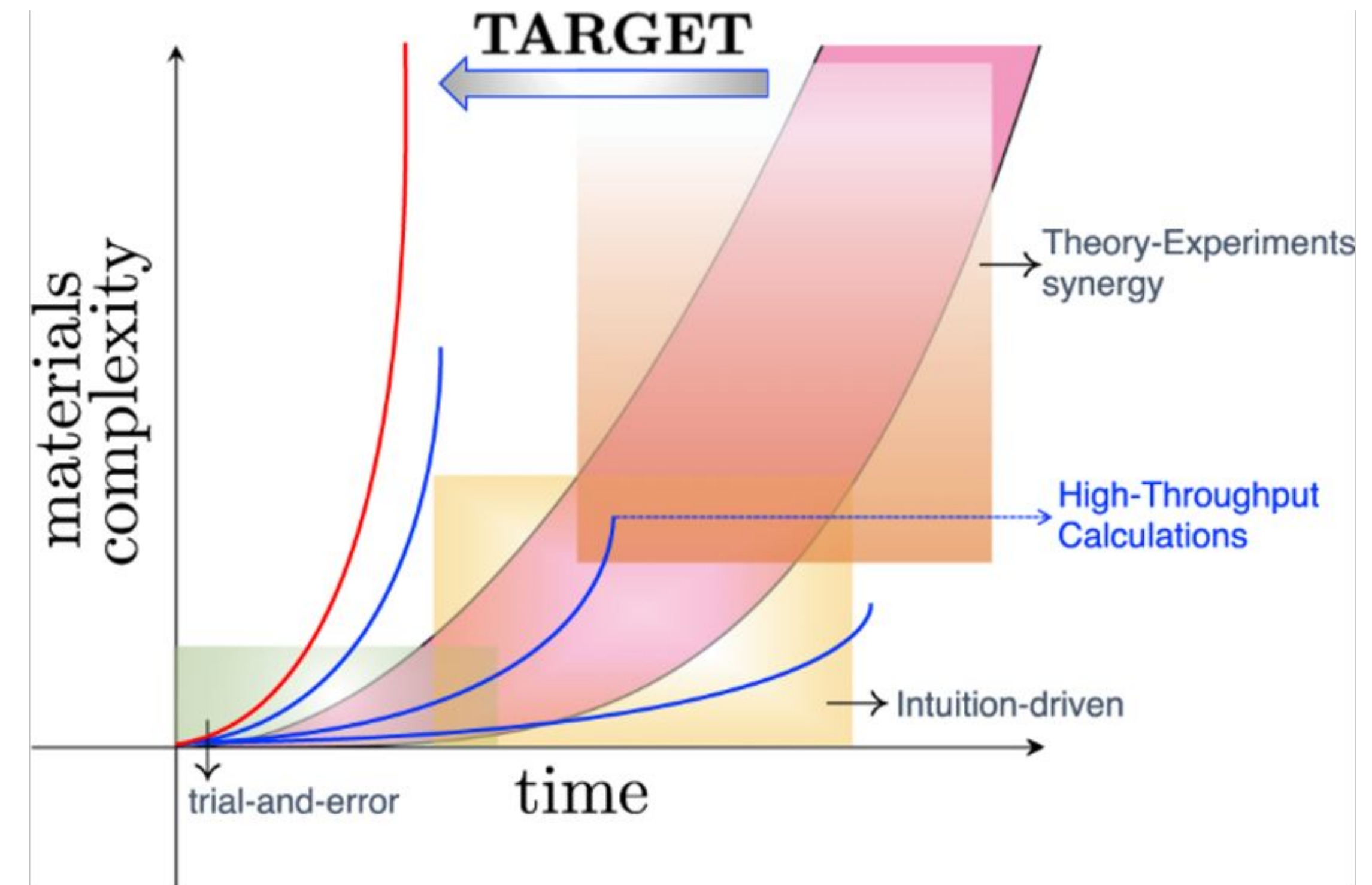




Glass Discovery

Motivation

- Glasses: back-bone of society
- Glass discovery: empirical, uneconomical with a ***design-to-deploy period of 20-30 years***
- UN sustainable development goals: 10/17 can potentially be addressed through glass innovation: IYG
- Areas: energy, infrastructure, healthcare, CO₂ emission, space sector



Lookman et al., npj Comp. Mat., 2019

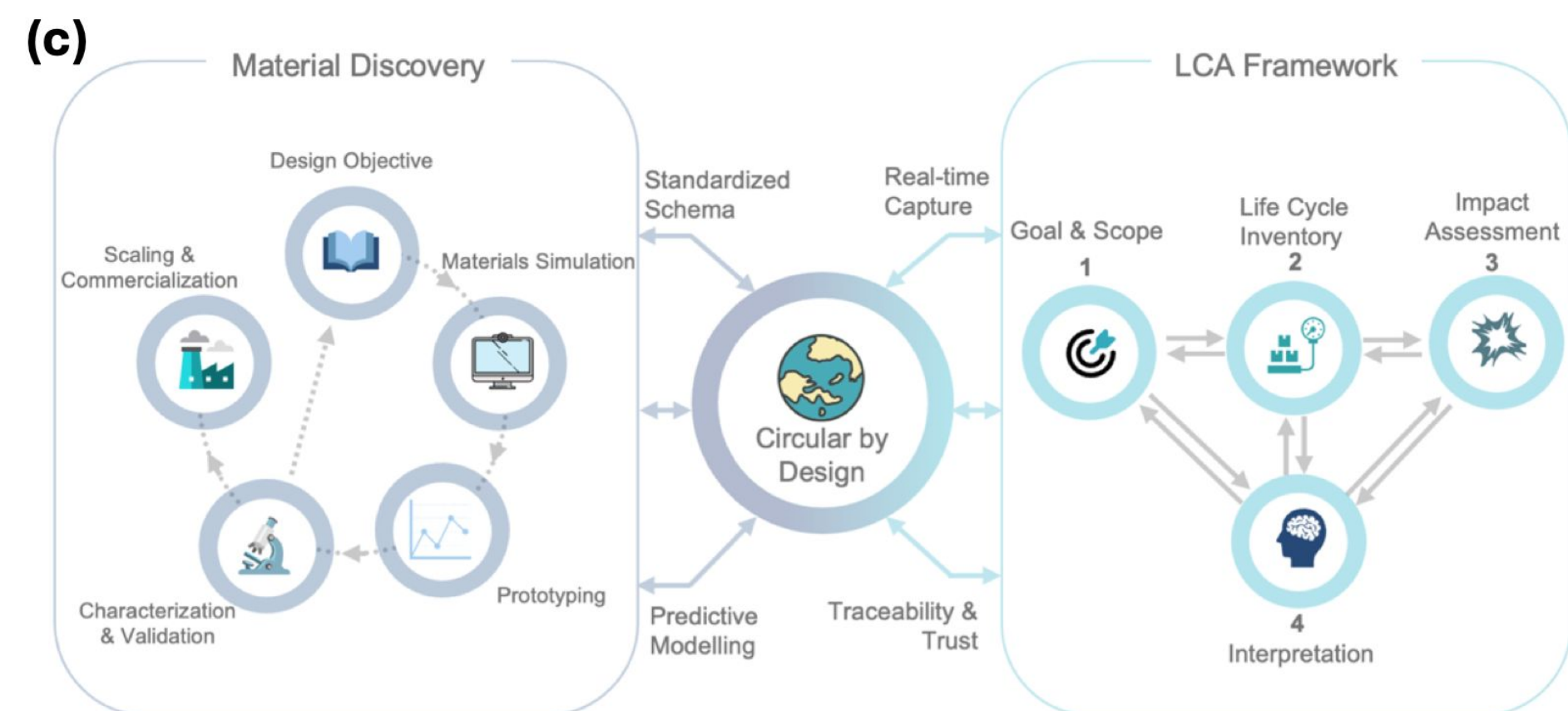
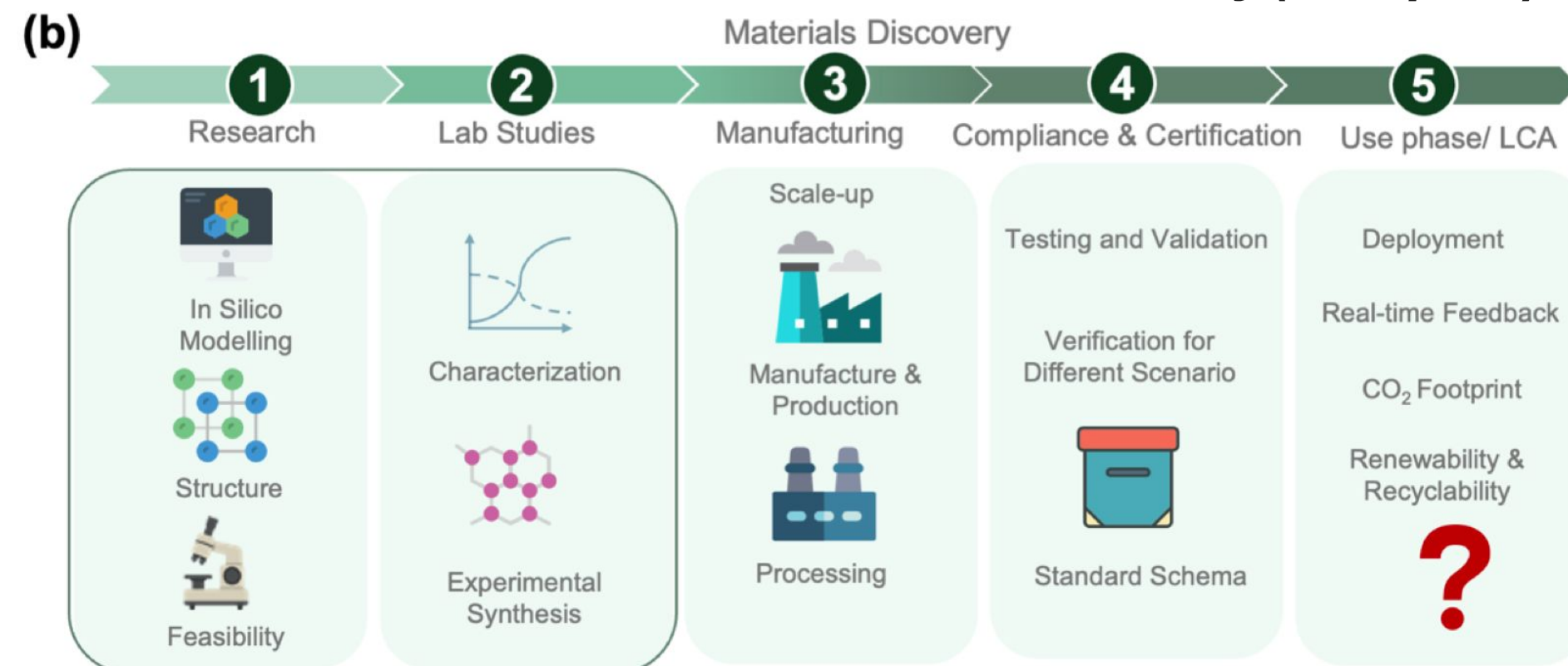
Alternate approach: Data-driven Modeling

AI Scientist for Glass Discovery

Sustainable Materials Discovery in the Era of Artificial Intelligence

Sajid Mannan¹, Rupert J. Myers², Rohit Batra³, Rocio Mercado⁴, Lothar Wondraczek^{5,6}, N. M. Anoop Krishnan^{1,7,*}

Nature Sustainability (accepted)



Overtrained Language Models Are Harder to Fine-Tune

Jacob Mitchell Springer¹ Sachin Goyal¹ Kaiyue Wen² Tanishq Kumar³ Xiang Yue¹
Sadhika Malladi⁴ Graham Neubig¹ Aditi Raghunathan¹

Kausik Hira¹, Mohd Zaki², Mausam^{1,3,#}, N. M. Anoop Krishnan^{1,2,#}

¹Yardi School of Artificial Intelligence, Indian Institute of Technology Delhi

²Department of Civil Engineering, Indian Institute of Technology Delhi

³Department of Computer Science and Engineering, Indian Institute of Technology Delhi

#Corresponding authors: {mausam, krishnan}@iitd.ac.in

Informatic
Machine intelligence

Industrial Scale
Manufacturing

AI-driven process
optimization at
industrial scale

nature computational science

[View all journals](#)

[Saved research](#)

[Explore content](#) ▾

[About the journal](#) ▾

[Publish with us](#) ▾

[nature](#) > [nature computational science](#) > [resources](#) > [article](#)

Resource | Published: 14 July 2026

UniFFBench: evaluating universal machine learning force fields against experimental measurements

[Sajid Mannan](#), [Vaibhav Bihani](#), [Carmelo Gonzales](#), [Kin Long Kelvin Lee](#), [Nitya Nand Gosvami](#), [Sayan Ranu](#),
[Santiago Miret](#) ✉ & [N. M. Anoop Krishnan](#) ✉

Nature Computational Science **6**, 894–906 (2026) | [Cite this article](#)

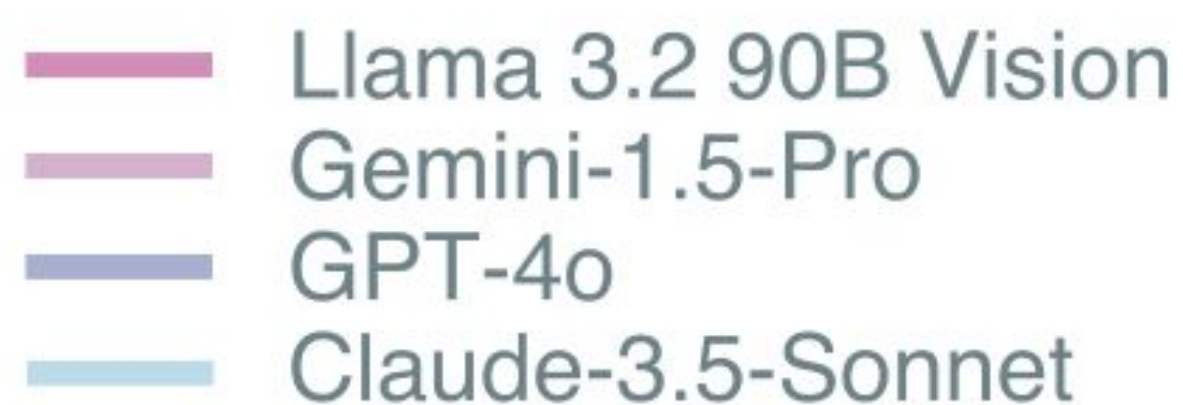
[Save article](#)

2038 Accesses | 1 Citations | 6 Altmetric | [Metrics](#)

MaCBench

Benchmarking multimodal models

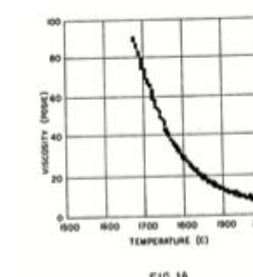
Models



Example questions

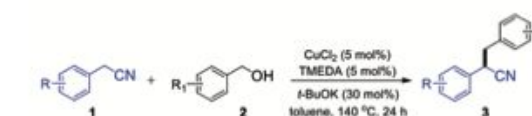
Tables
How many material compositions are present in the table?

Plots
What is the viscosity (in poise) of the fiber at a temperature of 2073 K?



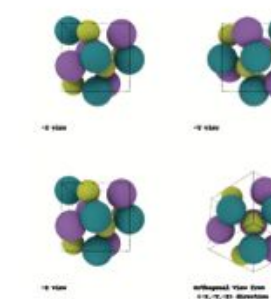
Organic reactions and molecules

Which solvent was used in the reaction?



In silico experiments

What is the volume of the unit cell of this crystal in Å³?

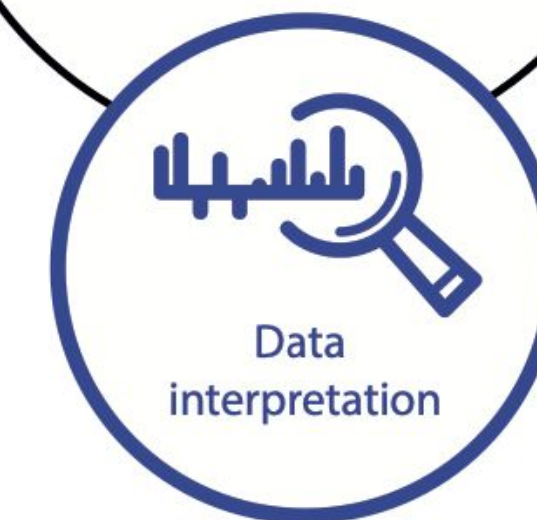


Laboratory experiments

The chemist wants to run a volatile chemical reaction overnight. What statements about the picture are true?

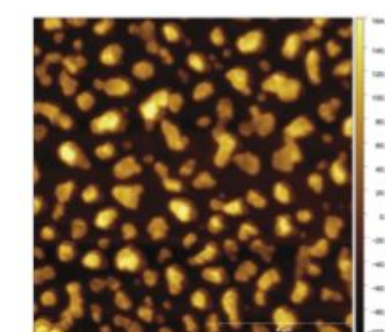


MaCBench



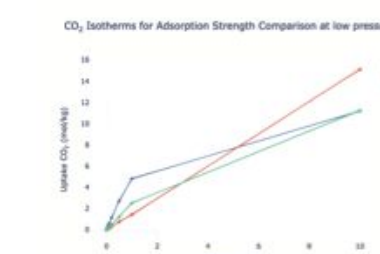
AFM characterization

What is the maximum height of the gold nanoislands in nanometers?



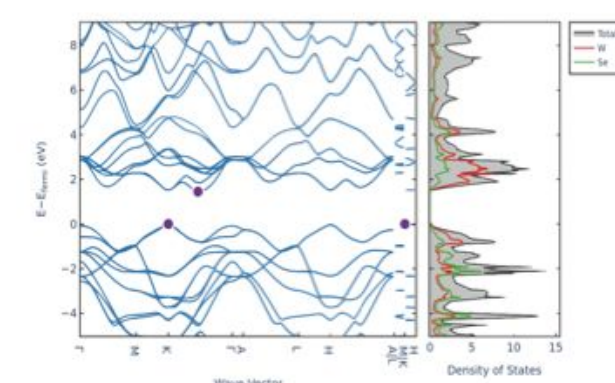
MOF characterization

What is the correct ascending order of adsorption strength at low pressure (0.01 bar) in mol/kg for CO₂ in the given MOF isotherms simulated at 298 K?



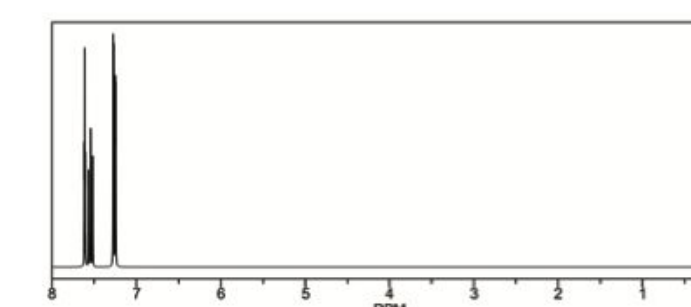
Bandgap analysis

Does the electronic structure in the following image show a bandgap?



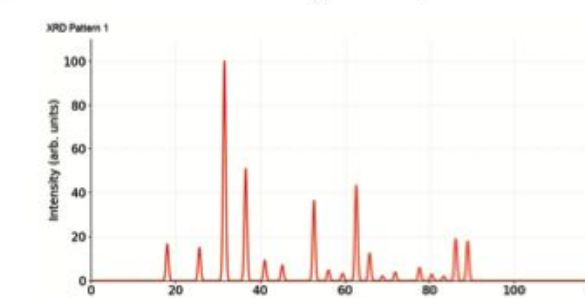
Molecule analysis

The measured molecule is an aromatic 6-membered ring with 2 chloride substituents. Based on the NMR spectrum, how are the substituents located to each other?



XRD pattern analysis

What is the crystal structure type of the material based on the following XRD pattern?





LLaMat

LLM for Materials

Pretraining



Peer-reviewed
MatSci Publications

Materials Science
Community Discourse

3.051%
0.019%
2.499%

Crystallographic
Information Files

RedPajama (subset)

Publications
Redpajama
MatSci Comm.
CIF

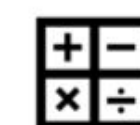
LLaMat

Instruction Finetuning

OpenOrca



MathQA



MatBookQA



MatSciNLP



MatSciInstruct



MaSQA



LLaMat-Chat

Materials IE co-pilot



MatNLP
MOF

Doping
General



	0	1	2	3	4
0 Sample	Li2O	Na2O	Al2O3	SiO2	
1 C1	1	0	1	4	
2 C2	0.8	0.2	1	4	
3 C3	0.6	0.4	1	4	
4 C4	0.5	0.5	1	4	



IE from tables



IE from text

Finetuning for Crystals



Syntactic
Instructions

What is the lattice
parameters of...?
What is the volume
of the unit cell?

Semantic
Instructions

Which atoms are
present in the
crystal?

What is the
coordinate of...?

Which element is
present with ...?

```
_chemical_formula_structural
'Mg2 (Si2 O6) '
_chemical_name_mineral
Enstatite
_symmetry_cell_setting
orthorhombic
_cell_angle_alpha
90
_cell_angle_beta
90
_cell_angle_gamma
90
_cell_formula_units_Z
8
_cell_length_a
18.25099(400)
_cell_length_b
8.814(1)
_cell_length_c
5.181(1)
...
```

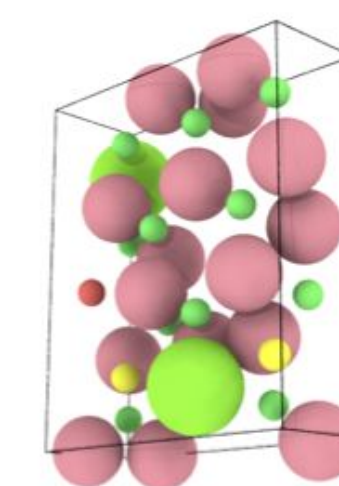
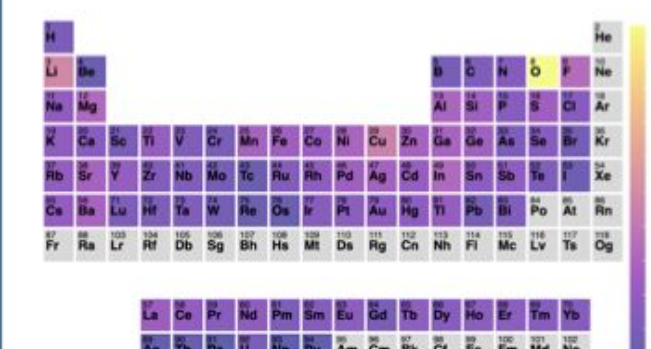


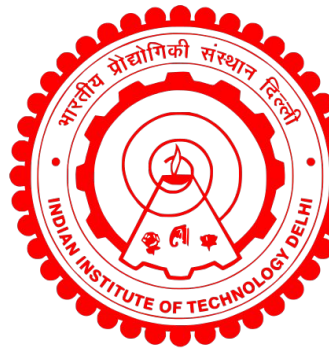
LLaMat-CIF

Crystal Generator



Generate a stable
crystal based on...

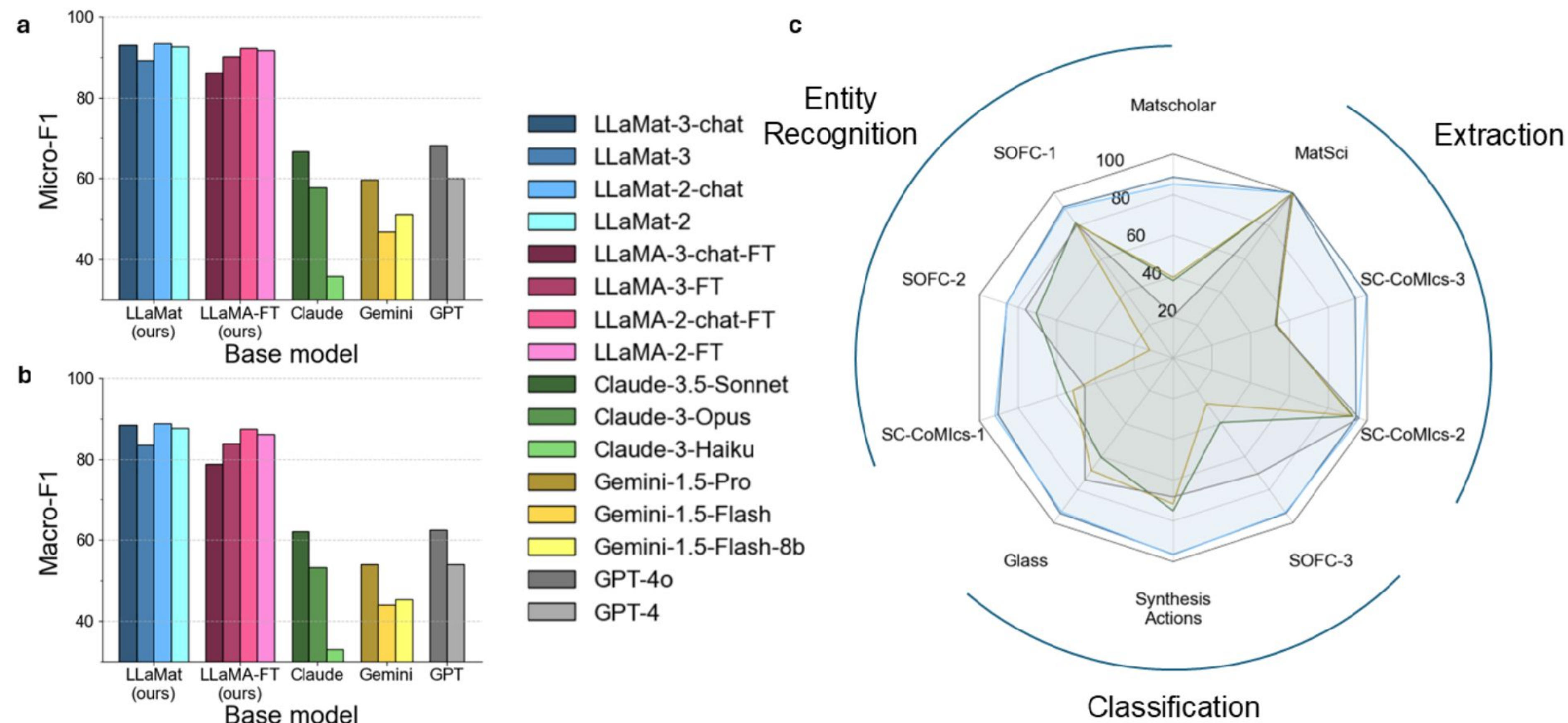




LLaMat

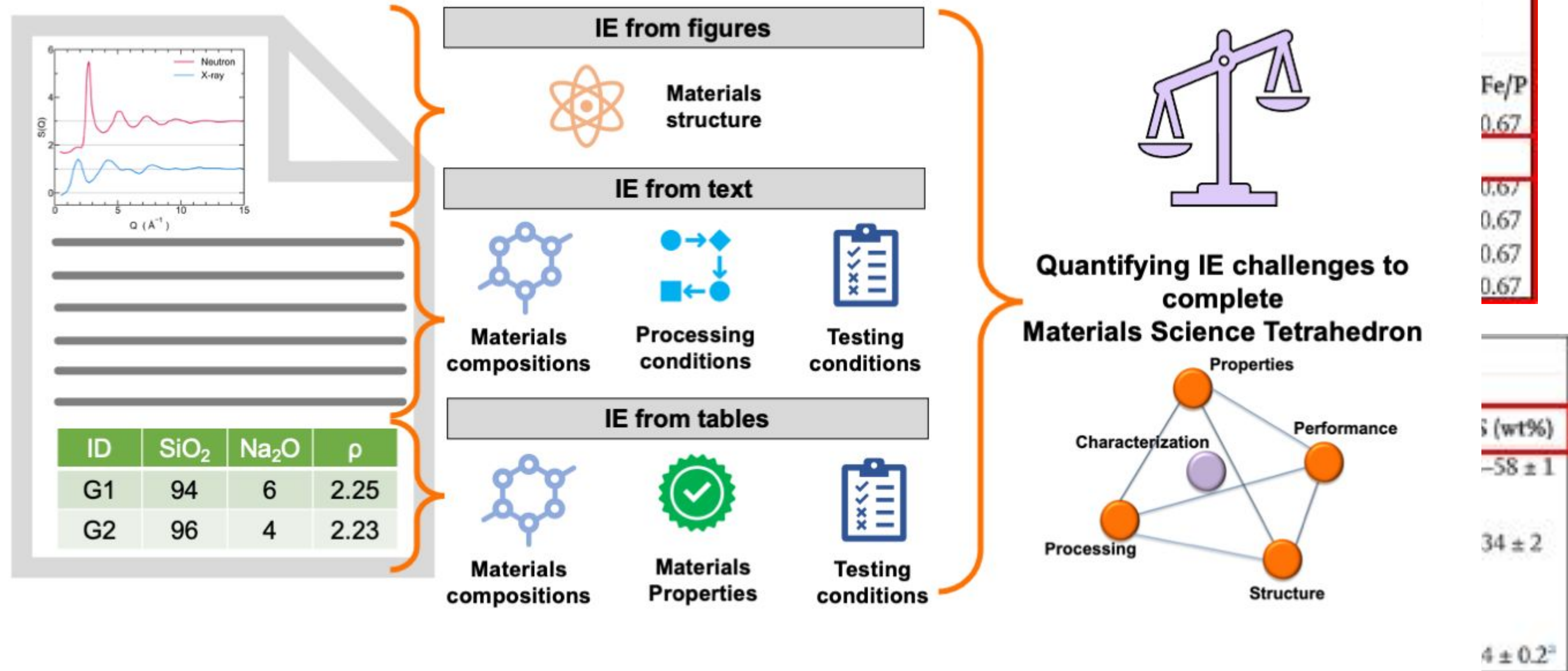
First Large Language Model for Materials

LLaMat
significantly
outperforms
LLMs on
MatSci tasks!



Information Extraction from Tables

Challenges





Information Extraction from Tables

Challenges

6. **Single-cell compositions:** Entire composition in one table cell (some coefficients may be missing)
7. **Percentages not summing to 100:** Re-normalization is needed. Common case is doping
8. **Percentage contributions as variables:** Like x, y, etc.
Contributions can have arithmetic expressions
9. **Partial Info Tables:** Percentages of only some constituents reported in table. Infer remaining composition from table caption or publication text
10. **Other corner cases**

Gupta et al., ACL, 2023

Composition	$\log \sigma_{298} (\text{S cm}^{-1})$
$(80\text{GeS}_2-20\text{Ga}_2\text{S}_3)_{90}-(\text{LiI})_{10}$	-6.34 (5)
$(80\text{GeS}_2-20\text{Ga}_2\text{S}_3)_{90}-(\text{NaCl})_{10}$	-6.92 (5)

Base composition(mol%):75TeO ₂ -20ZnO-5Na ₂ O			
Glass	Tm ₂ O ₃ (wt%)	Yb ₂ O ₃ (wt%)	Ho ₂ O ₃ (wt%)
Undoped TZN	–	–	–
TT05	0.5	–	–
TTY20	1	2	–
TTH05	1	–	0.5
TTYH	1	5	0.15

	$(\text{GeSe}_2)_{1-x}(\text{GeBr}_4)_x$		$(\text{GeSe}_2)_{1-x}\text{Br}_x$		
x	0.079	0.167	0.265	0.250	0.429
p _{Se} (%)	94.1	86.0	76.5	84.1	78.0

Glass composition in the paper: $20\text{La}_2\text{S}_3 - (80 - x)\text{Ga}_2\text{S}_3 - x\text{CsCl}$		
Glass	CsCl (mol%)	n (1.5 μm)
GLSC10	10	2.253 ± 0.001
GLSC20	20	2.265 ± 0.001

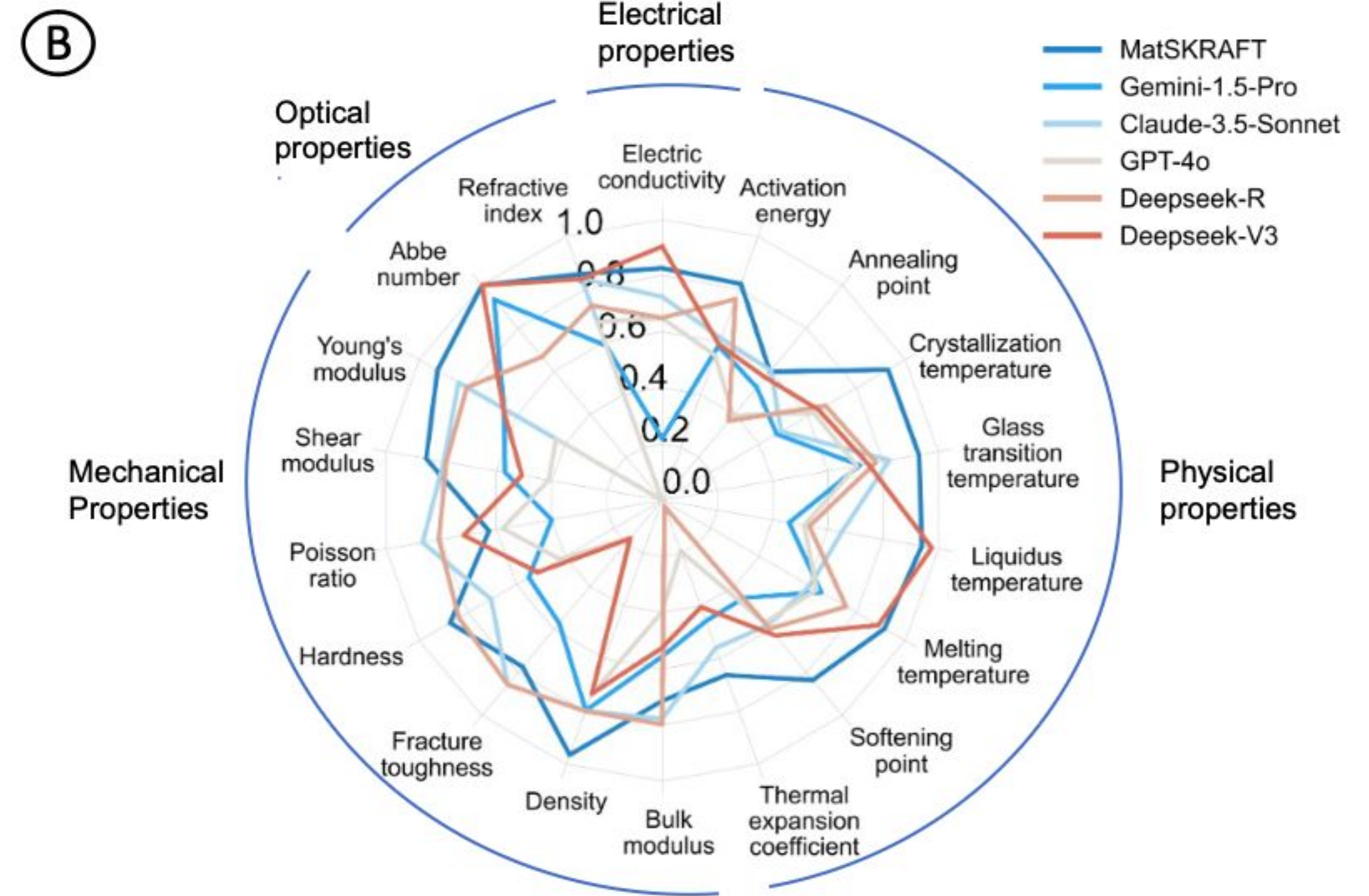
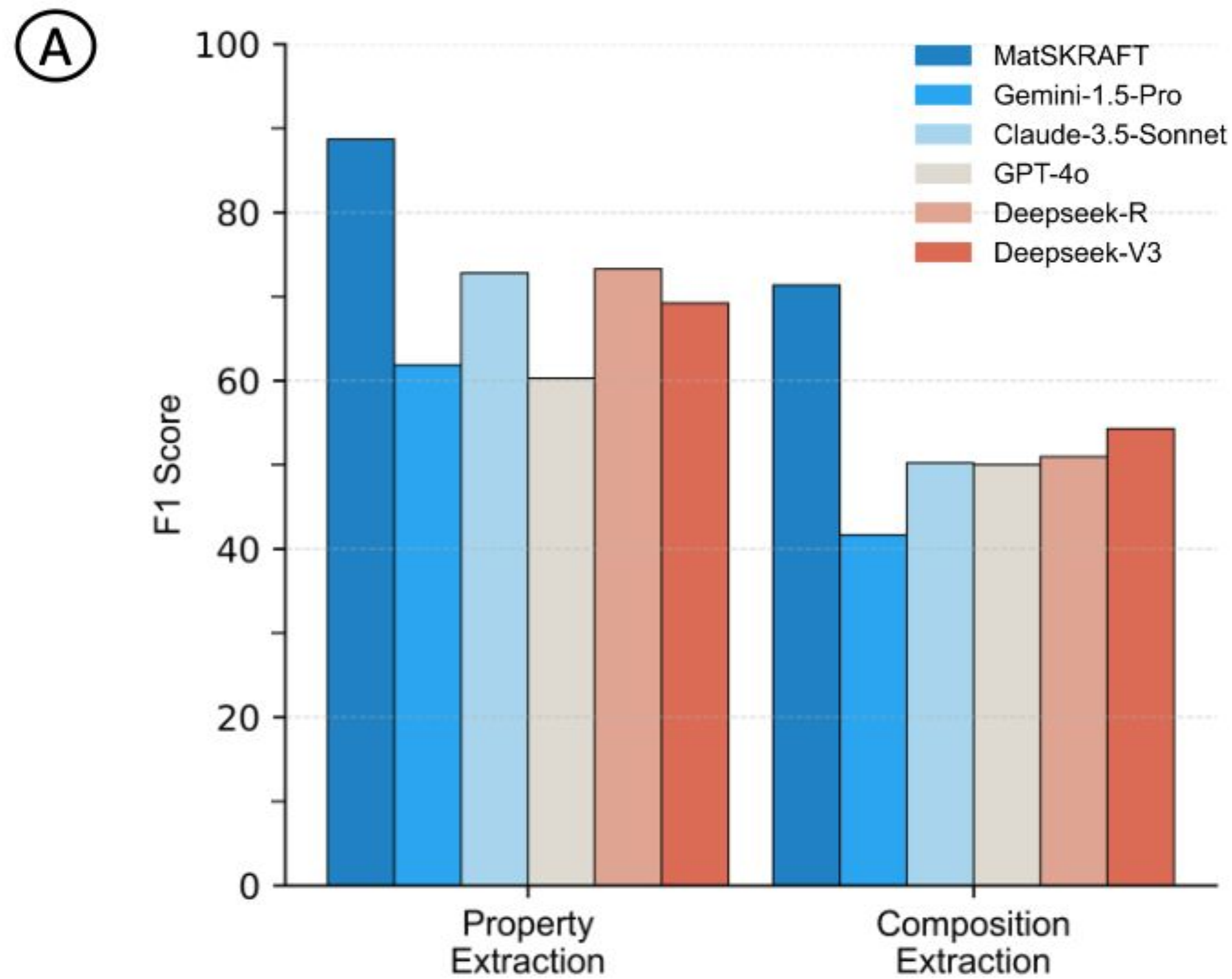
MatSKRAFT

Composition extraction from tables

1. Property extraction with units
2. Composition extraction
3. Linking them
4. Applications

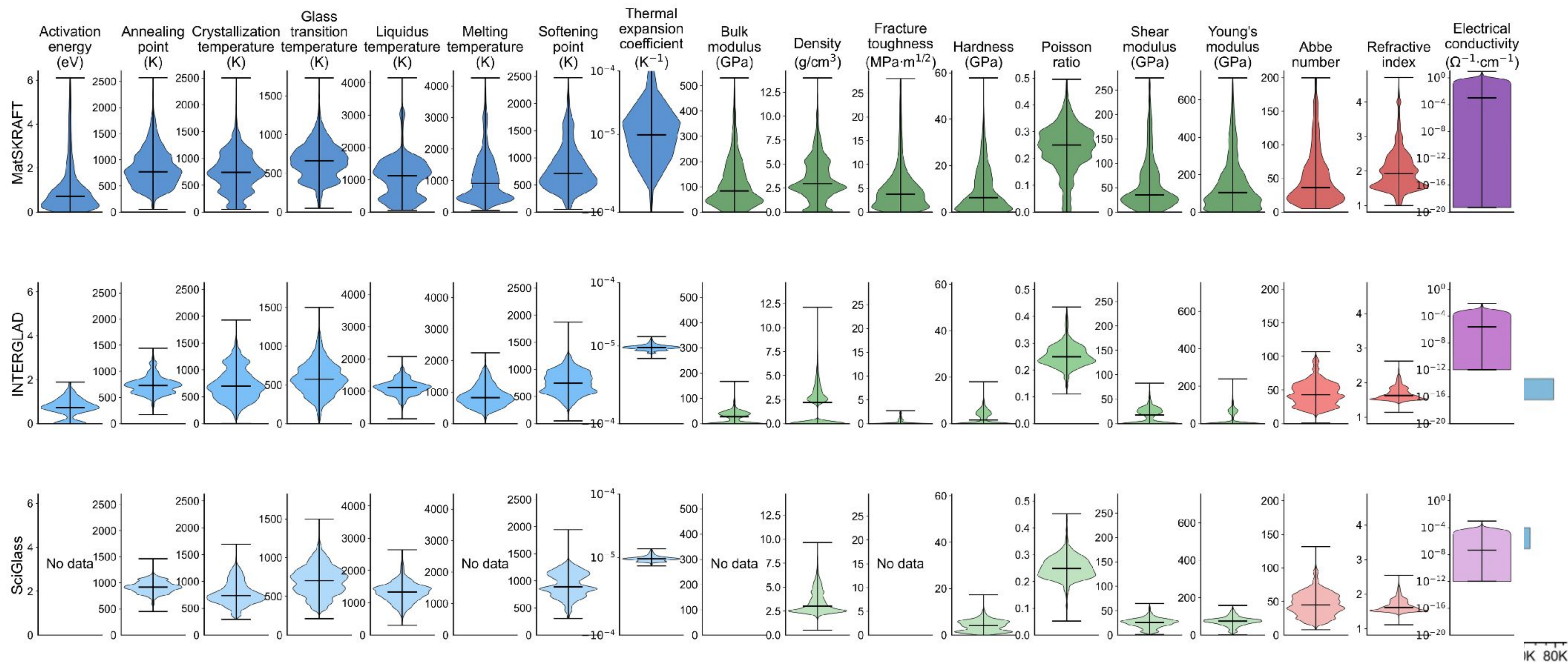
MatSKRAFT

Composition and property extraction from tables

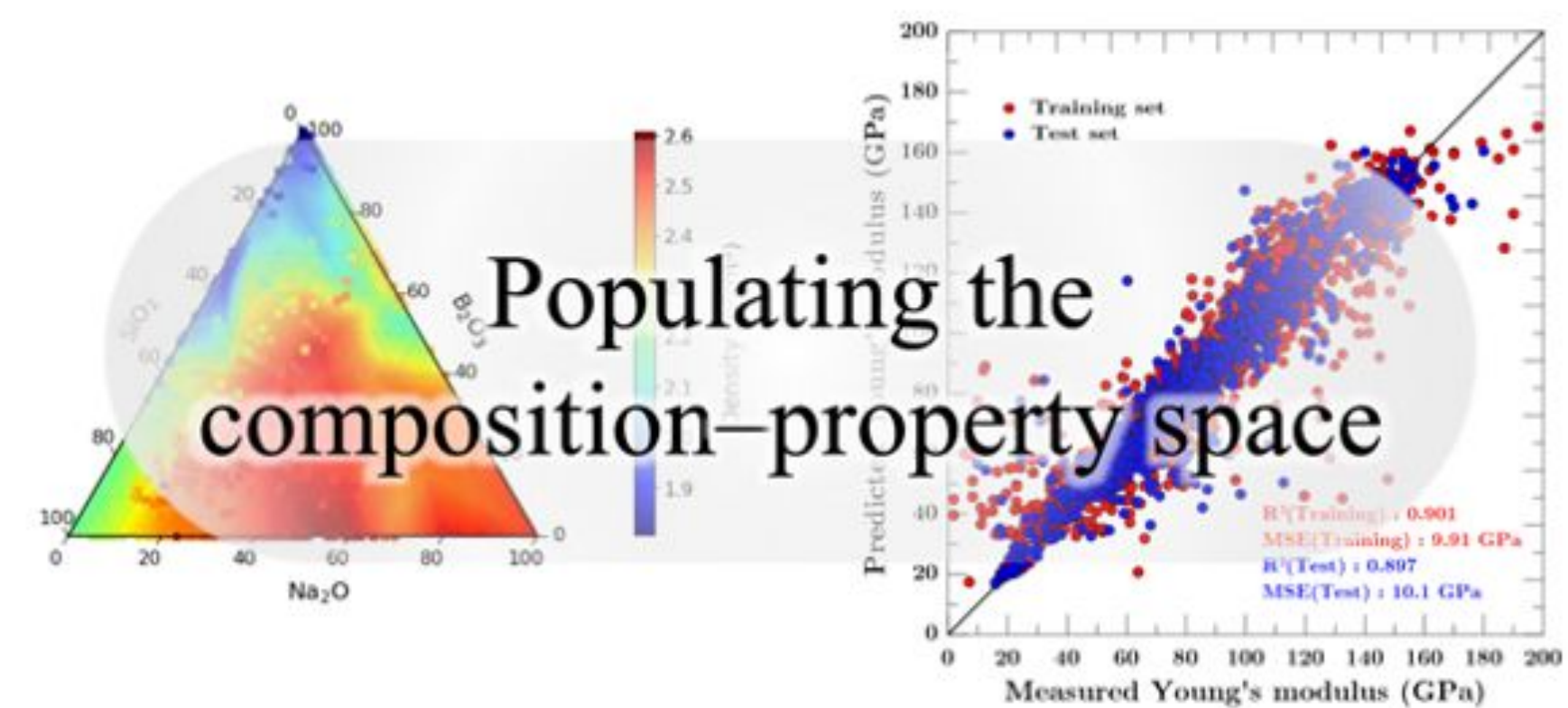
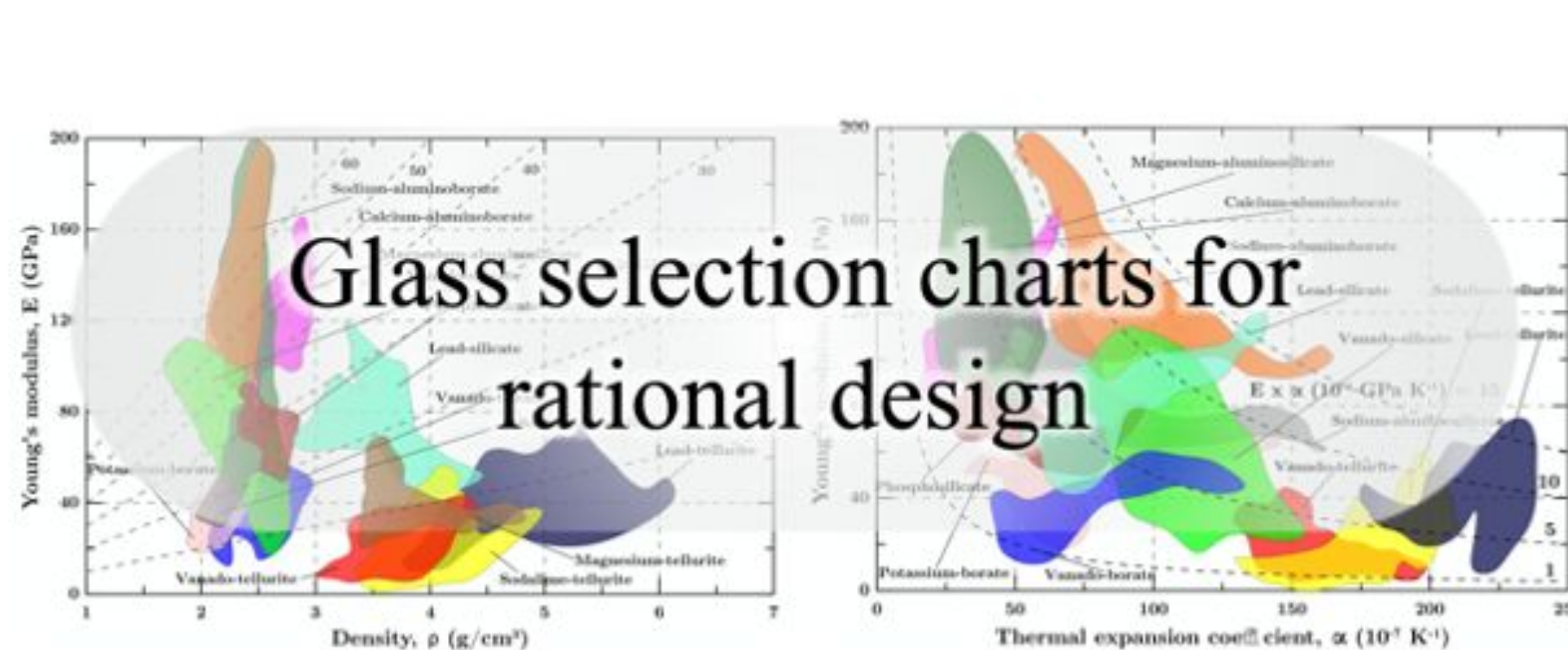
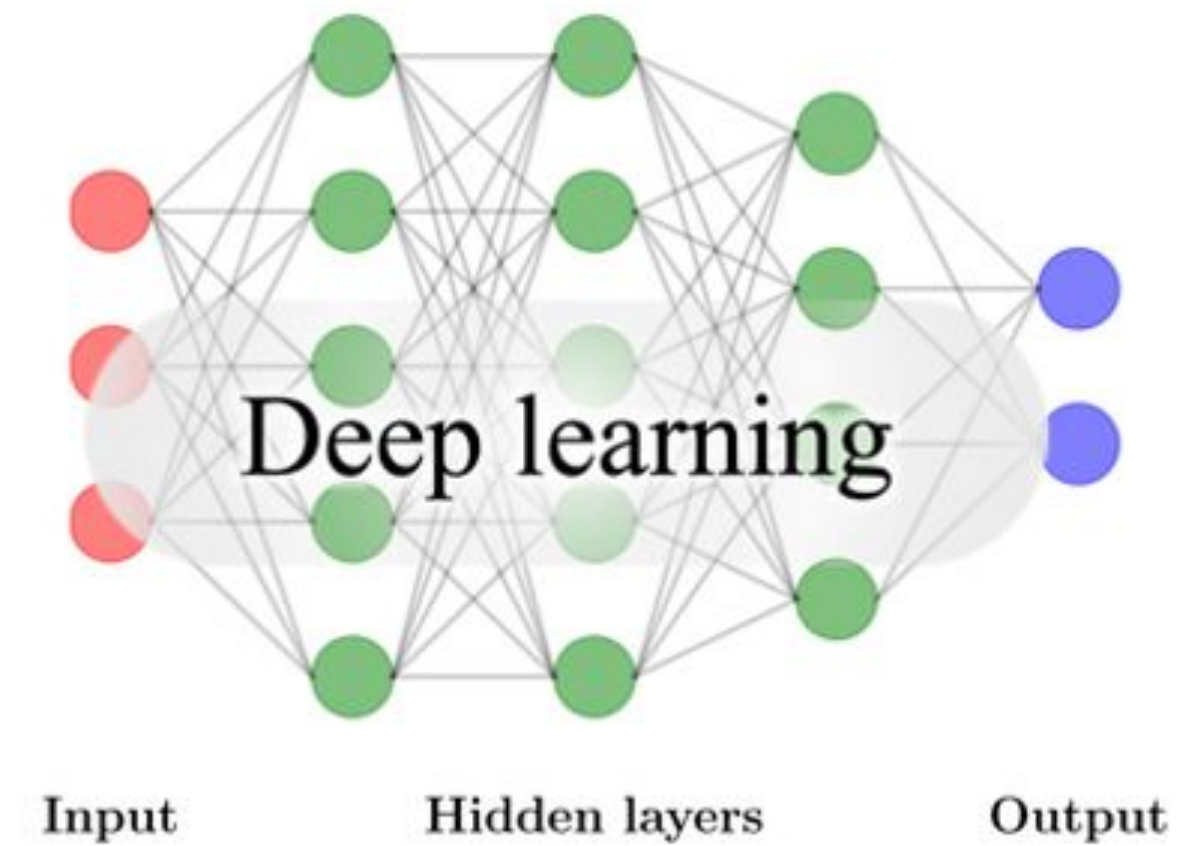
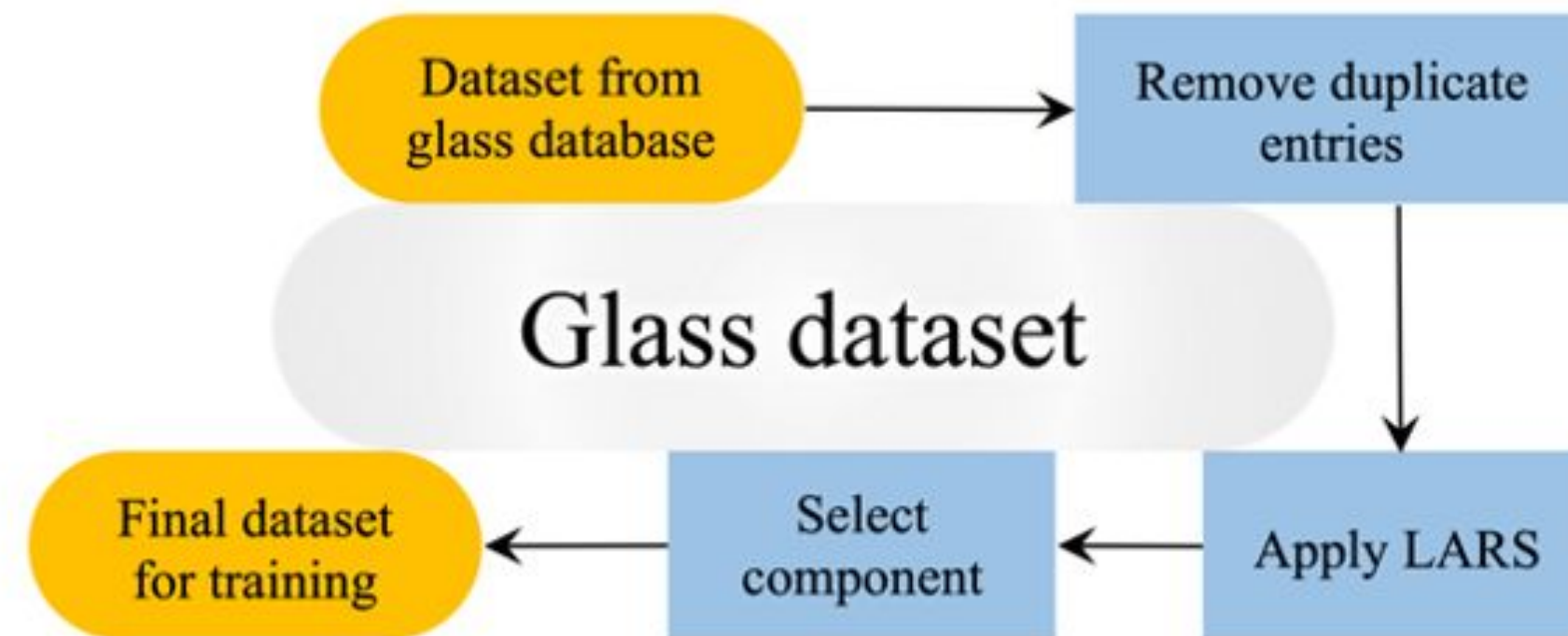


MatSKRAFT

Extracted dataset



Data-driven modeling in glasses



Dataset visualization

1 H Hydrogen																	2 He Helium				
3 Li Lithium	4 Be Beryllium															5 B Boron	6 C Carbon	7 N Nitrogen	8 O Oxygen	9 F Fluorine	10 Ne Neon
11 Na Sodium	12 Mg Magnesium															13 Al Aluminium	14 Si Silicon	15 P Phosphorus	16 S Sulfur	17 Cl Chlorine	18 Ar Argon
19 K Potassium	20 Ca Calcium	21 Sc Scandium	22 Ti Titanium	23 V Vanadium	24 Cr Chromium	25 Mn Manganese	26 Fe Iron	27 Co Cobalt	28 Ni Nickel	29 Cu Copper	30 Zn Zinc	31 Ga Gallium	32 Ge Germanium	33 As Arsenic	34 Se Selenium	35 Br Bromine	36 Kr Krypton				
37 Rb Rubidium	38 Sr Strontium	39 Y Yttrium	40 Zr Zirconium	41 Nb Niobium	42 Mo Molybdenum	43 Tc Technetium	44 Ru Ruthenium	45 Rh Rhodium	46 Pd Palladium	47 Ag Silver	48 Cd Cadmium	49 In Indium	50 Sn Tin	51 Sb Antimony	52 Te Tellurium	53 I Iodine	54 Xe Xenon				
55 Cs Caesium	56 Ba Barium	57 La Lanthanum	72 Hf Hafnium	73 Ta Tantalum	74 W Tungsten	75 Re Rhenium	76 Os Osmium	77 Ir Iridium	78 Pt Platinum	79 Au Gold	80 Hg Mercury	81 Tl Thallium	82 Pb Lead	83 Bi Bismuth	84 Po Polonium	85 At Astatine	86 Rn Radon				
87 Fr Francium	88 Ra Radium	89 Ac Actinium	104 Rf Rutherfordium	105 Db Dubnium	106 Sg Seaborgium	107 Bh Bohrium	108 Hs Hassium	109 Mt Meitnerium	110 Ds Darmstadtium	111 Rg Roentgenium	112 Cn Copernicium	113 Nh Nihonium	114 Fl Flerovium	115 Mc Moscovium	116 Lv Livermorium	117 Ts Tennesine	118 Og Oganesson				
			58 Ce Cerium	59 Pr Praseodymium	60 Nd Neodymium	61 Pm Promethium	62 Sm Samarium	63 Eu Europium	64 Gd Gadolinium	65 Tb Terbium	66 Dy Dysprosium	67 Ho Holmium	68 Er Erbium	69 Tm Thulium	70 Yb Ytterbium	71 Lu Lutetium					
			90 Th Thorium	91 Pa Protactinium	92 U Uranium	93 Np Neptunium	94 Pu Plutonium	95 Am Americium	96 Cm Curium	97 Bk Berkelium	98 Cf Californium	99 Es Einsteinium	100 Fm Fermium	101 Md Mendelevium	102 No Nobelium	103 Lr Lawrencium					

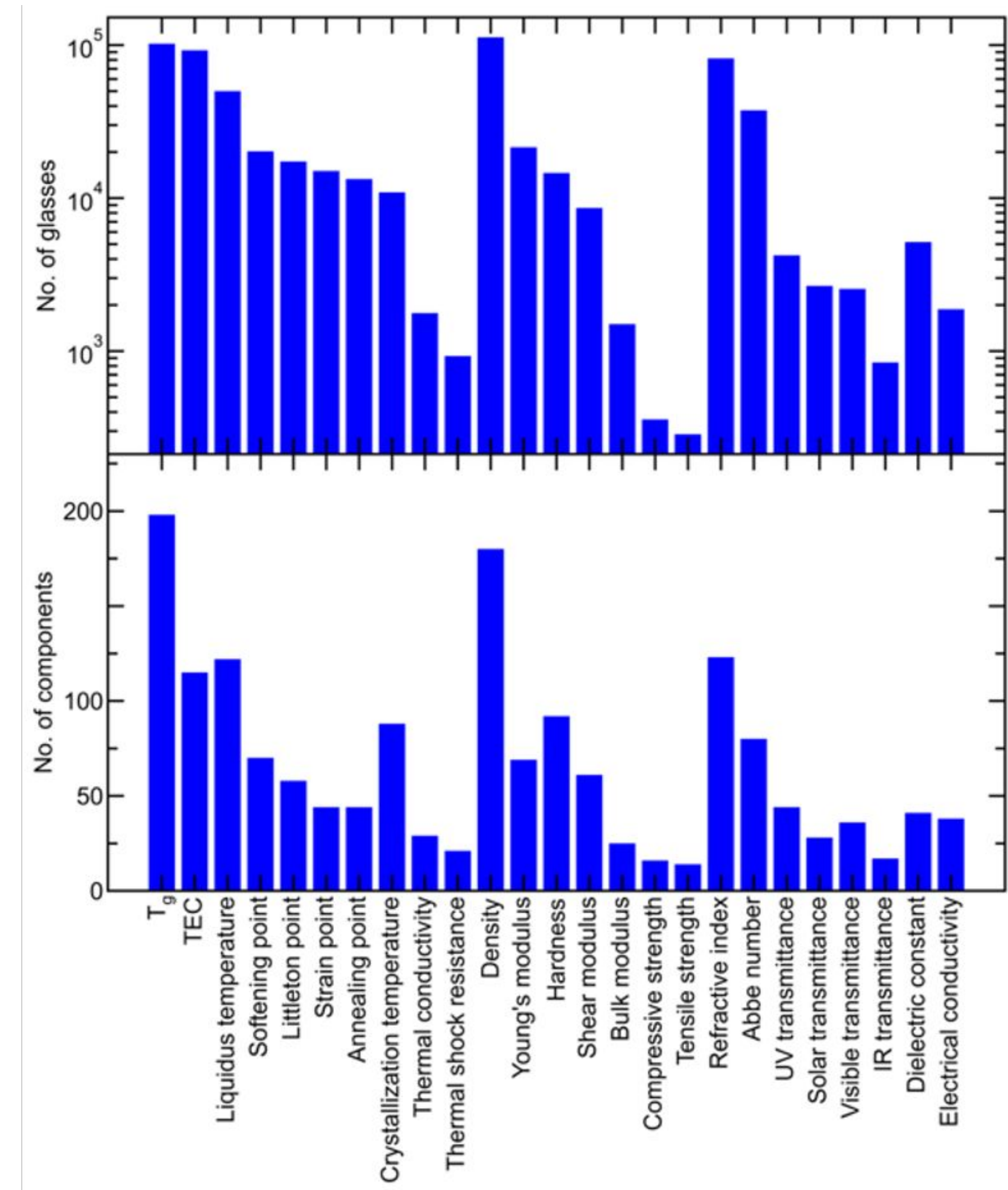
Network modifier

Network former

Intermediates

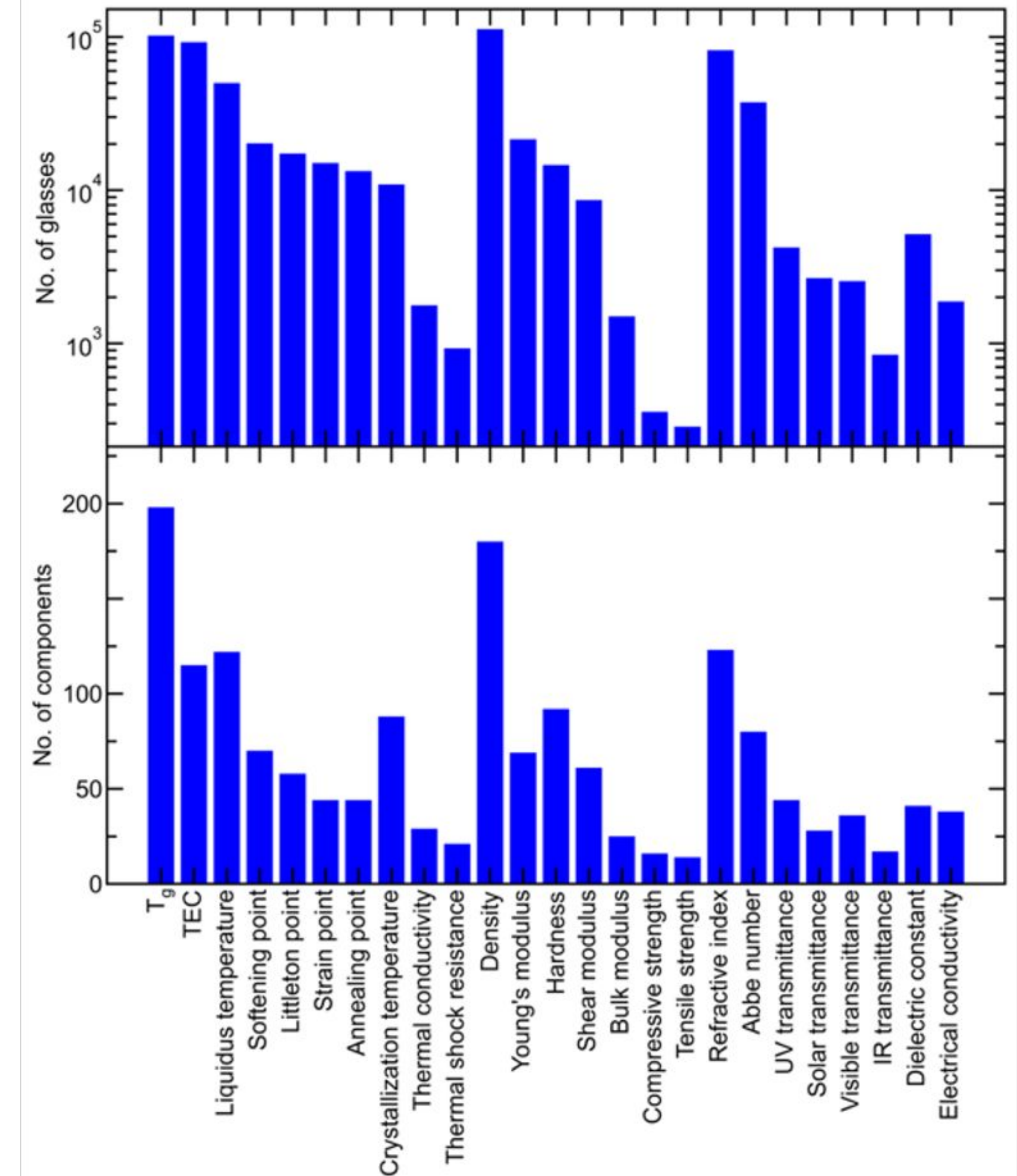
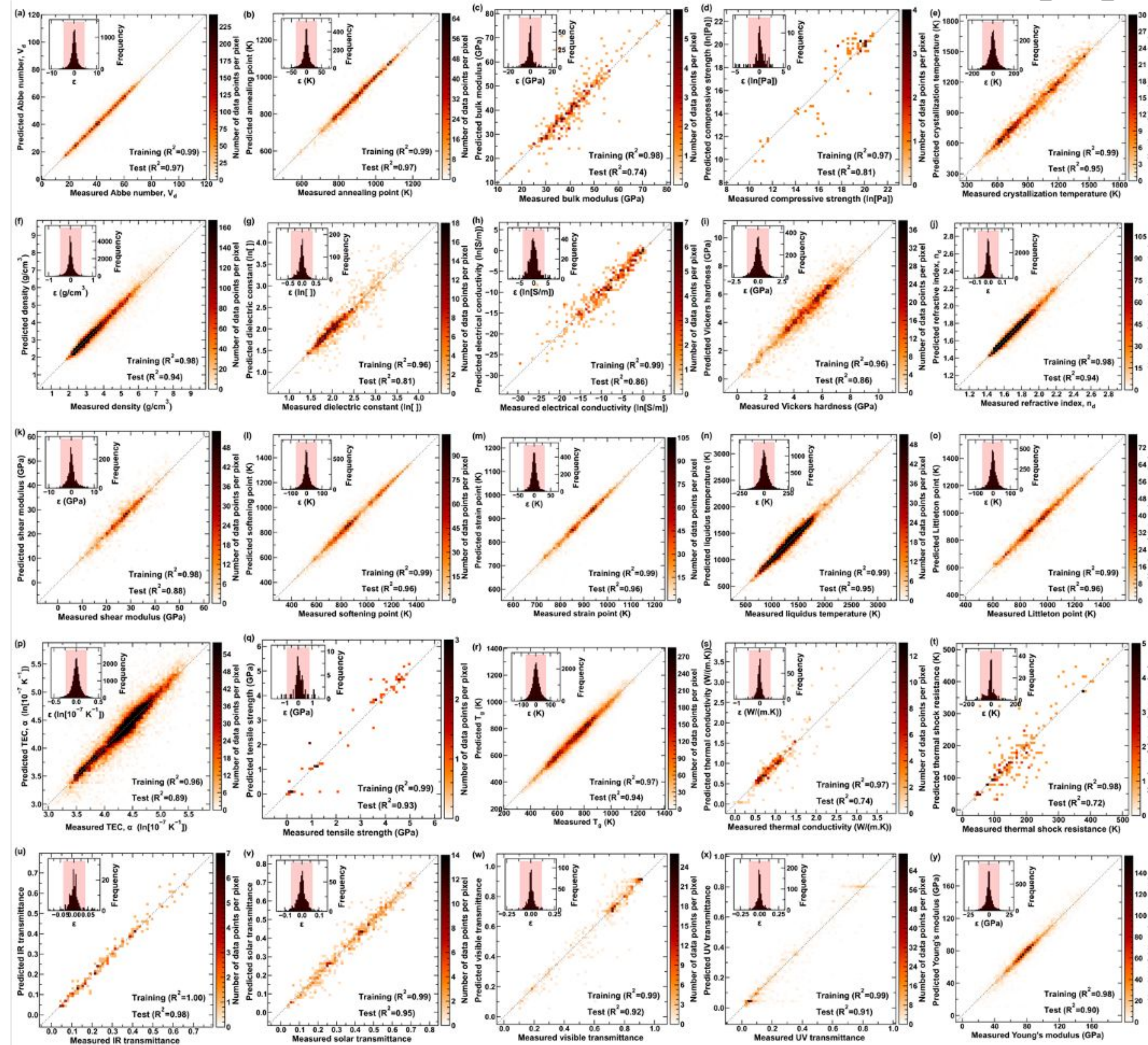
Others

Not present

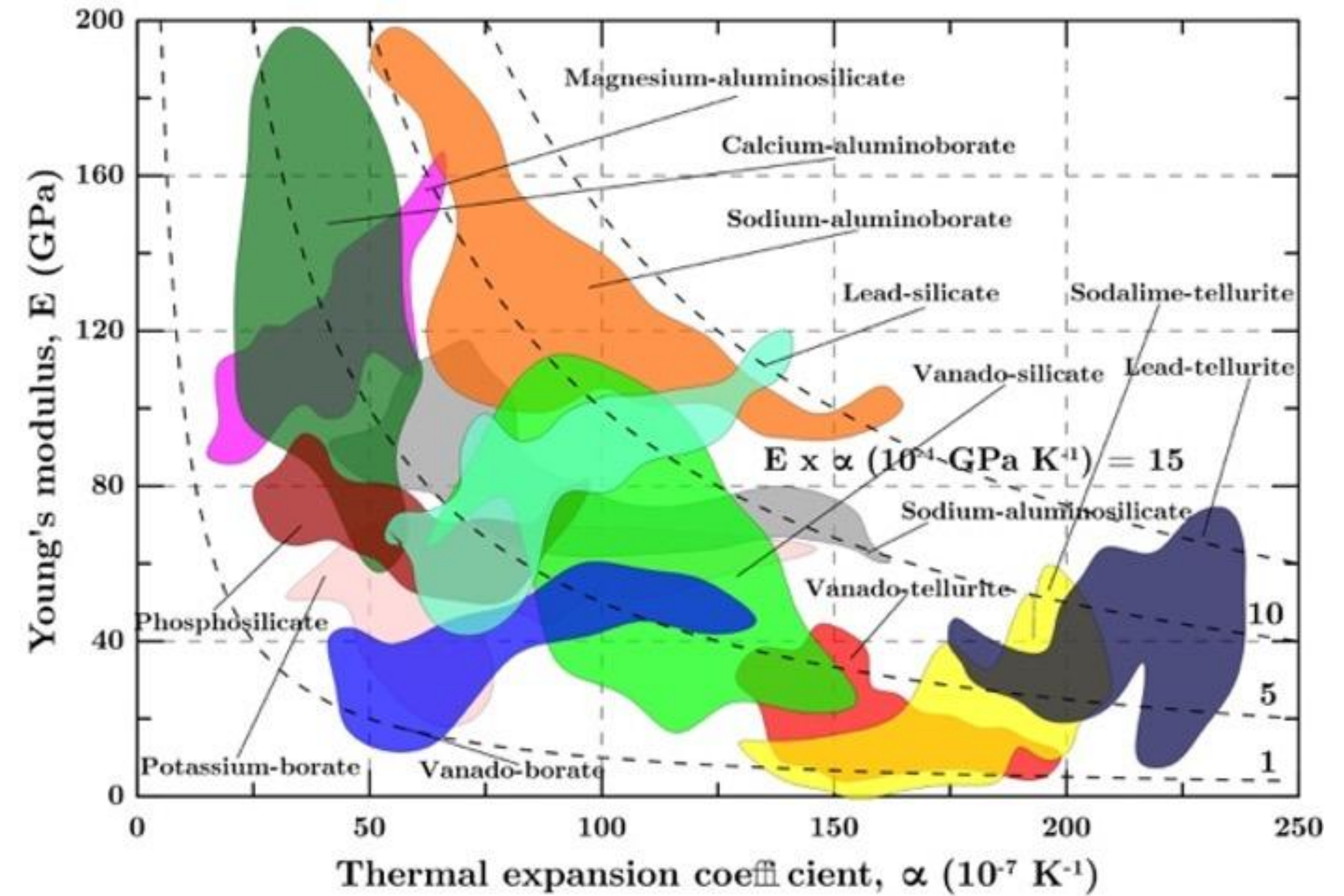
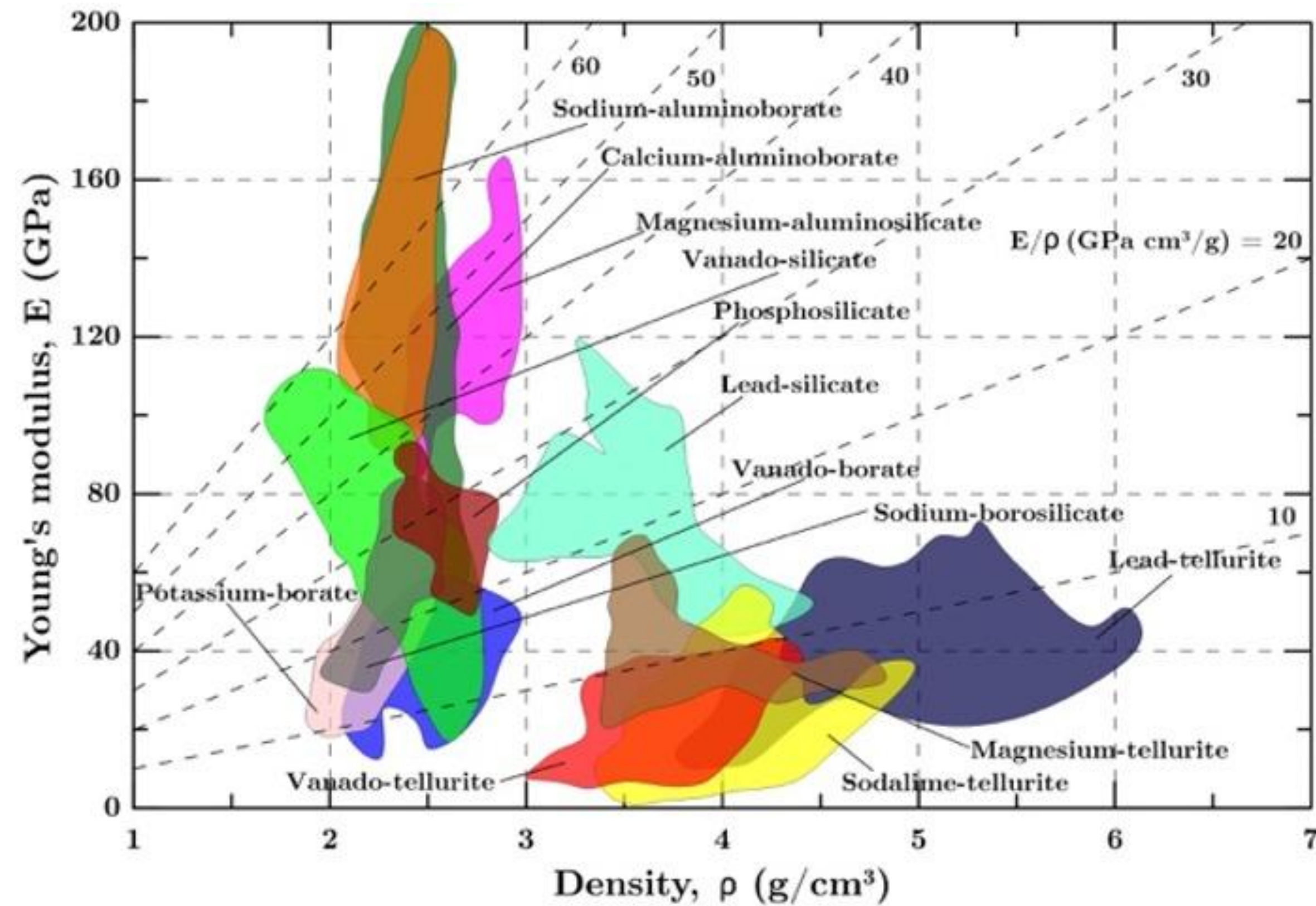


25 properties; 84 elements; 224 components; 450,000 glass compositions

Machine learning models



Material selection chart



Python for Glass Genomics

<https://pyggi.substantial.ai>



[Upcoming Events](#) [Store](#) [Donate](#) [LOG IN](#)

[Membership](#) [Meetings & Events](#) [IYOG](#) [Publications](#) [Professional Resources](#) [Learning Center](#) [Diversity & Inclusion](#) [Get Involved](#) [About](#) [🔍](#)

[Home](#) / [Ceramic Tech Today](#) / [Basic science](#) / [Deep learning provides deep help—researchers develop publicly available software for rational design of oxide glasses](#)

Deep learning provides deep help— researchers develop publicly available software for rational design of oxide glasses

- **ISRO: Glasses for space application**
- **BRNS: Glasses for nuclear waste immobilization**
- **Sisecam: Glasses for commercial applications**
- **Battery electrolyte glasses**

PyGGi Python for Glass Genomics



Python for Glass Genomics (PyGGi) is a package for accelerating innovations in the field of glasses and ceramics. It uses state-of-the-art machine learning algorithms and computational techniques to enable the users in designing novel glasses and understanding the composition-property relationships.

[Get Started](#)

[About PyGGi](#)

Overview



PyGGi Bank

Explore composition-property database of glasses collected from literature

4

Packages



PyGGi Seer

Predict glass properties as a function of the glass composition

9

Properties



PyGGi Zen

Discover new glass compositions based on target properties and compositional constraints

34

Oxide
Components

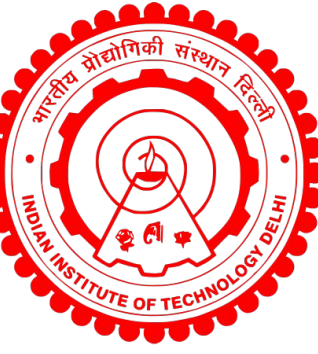


PyGGi Pictionary

Explore published works and images from glass literature

50000+

Glasses

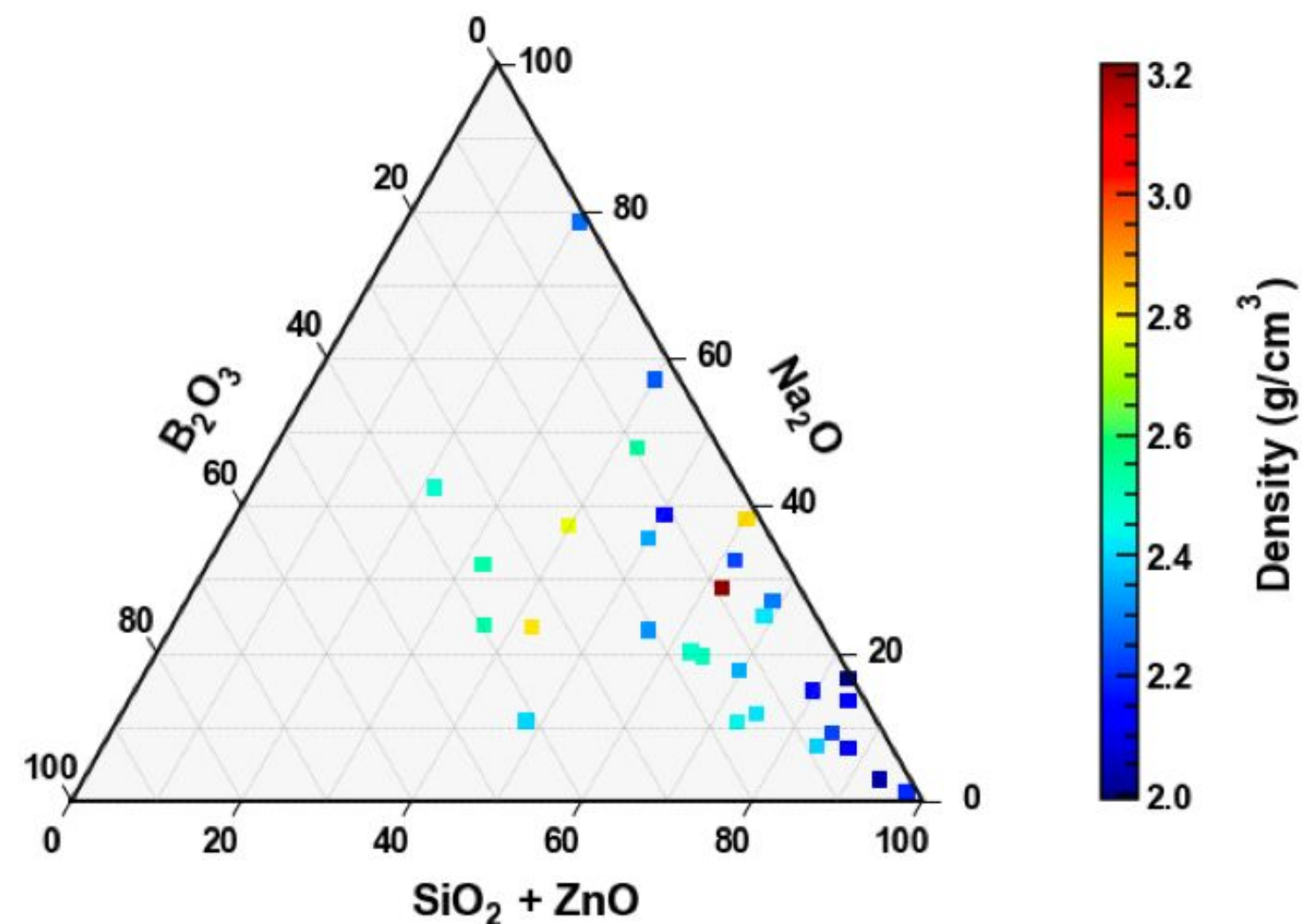


Case Studies

Constraints for nuclear waste glass

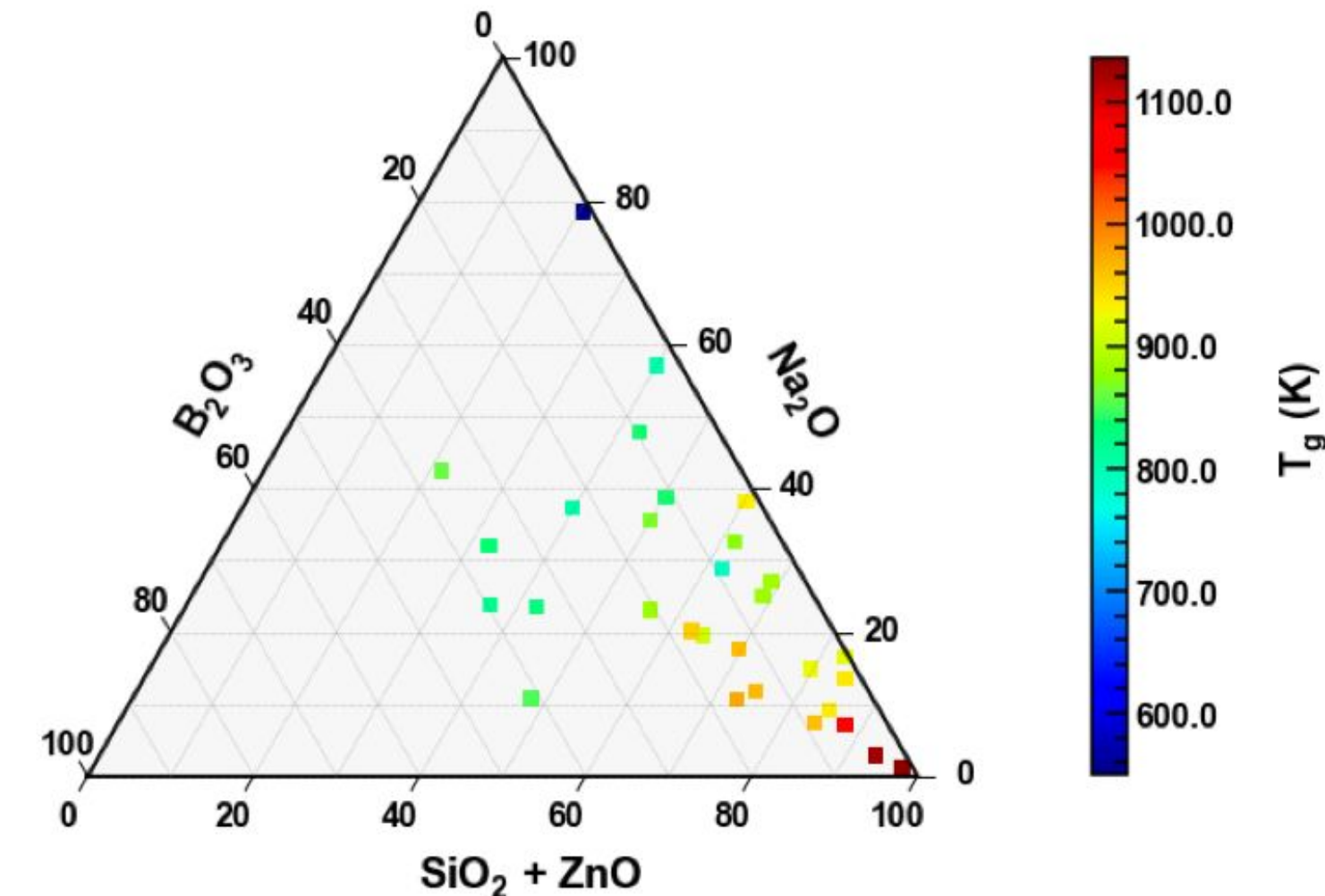
Composition constraints:

- a) $\text{SiO}_2 + \text{B}_2\text{O}_3 \sim 60\text{-}70\%$ (wt) is preferred with $\text{SiO}_2 > 45\%$ (wt) with $\text{B}_2\text{O}_3 > 10$ wt%
- b) $\text{Na}_2\text{O} > 15\%$ wt
- c) $\text{Na}_2\text{O}/\text{B}_2\text{O}_3 > 1$, preferably 1.25
- d) $\text{ZnO} \sim 2.5 - 10$ wt% with a never exceed to 15% wt%



Property constraints:

- a) T_g value $\Rightarrow 550$ $^{\circ}\text{C}$
- b) Density = $2\text{-}3$ g/cm^3
- c) Electrical conductivity = 50 S/cm
- d) Viscosity > 50 poise at pouring temperature
- e) Dissolution rate $< 10^{-5}$ to 10^{-6} $\text{g/cm}^2/\text{day}$)



Optimized glass compositions

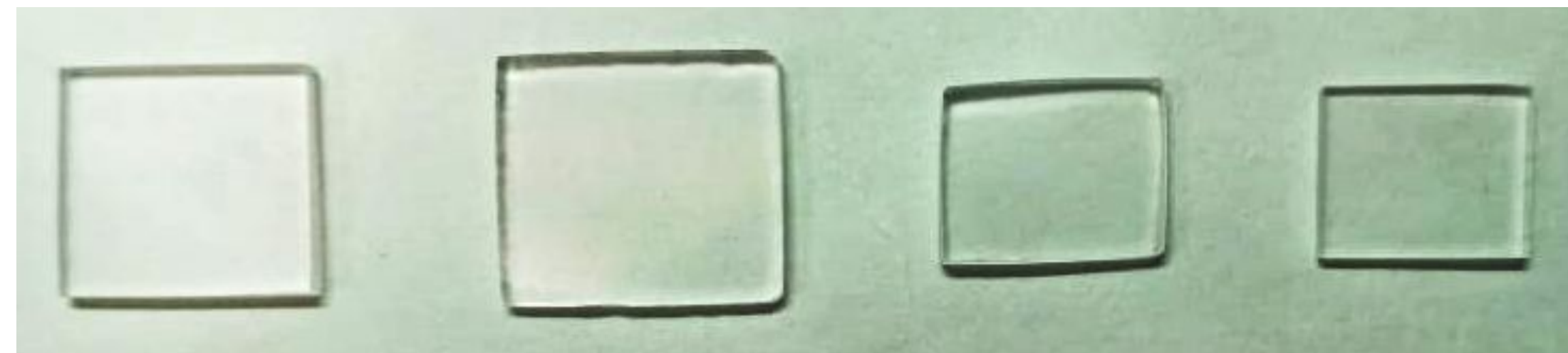
❖ Series of glasses obtained after compositional optimization using machine learned model

Na₂O (mol%)	SiO₂ (mol%)	B₂O₃ (mol%)	ZnO (mol%)	TEC (ln[10⁻⁷ K⁻¹)	Thermal Conductivity (W/(m.K))	Density (g/cm3)	T_g (K)	Viscosity (poise)
21	52	18	9	101.90	0.77	2.62	808.87	2.6831
24	52	18	6	105.57	0.78	2.59	785.74	3.0581
27	52	18	3	114.34	0.77	2.51	776.94	3.5894
30	52	18	0	121.98	0.75	2.52	779.55	1.4328

Glass preparation

Sample Name	Na ₂ O	SiO ₂	B ₂ O ₃	ZnO
NBSZ-9	21	52	18	9
NBSZ-6	24	52	18	6
NBSZ-3	27	52	18	3
NBSZ-0	30	52	18	0

Temperature (°C)	Time (h)
0-300	3
300	4
300-800	5
800	4
800-1200	2
1200	1



NBSZ-9

NBSZ-6

NBSZ-3

NBSZ-0

Comparison between ML and experiments

Na ₂ O(mol%)	SiO ₂ (mol %)	B ₂ O ₃ (mol%)	ZnO (mol %)	TEC (ln[10−7 K −1])	Thermal Cond uctivity (W/(m. K))	Density predicted (g/cm ³)	Density measured (g/cm ³)	Tg (K)	T _g measured (K)
21	52	18	9	101.90	0.77	2.62	2.61	809	803
24	52	18	6	105.57	0.78	2.59	2.60	786	806
27	52	18	3	114.34	0.77	2.51	2.56	777	793
30	52	18	0	121.98	0.75	2.52	2.52	780	785

Constraints for pharmaceutical glass (Sisecam)

Target

Target	Value
Softening Point	1058 K
No. of Iterations	50

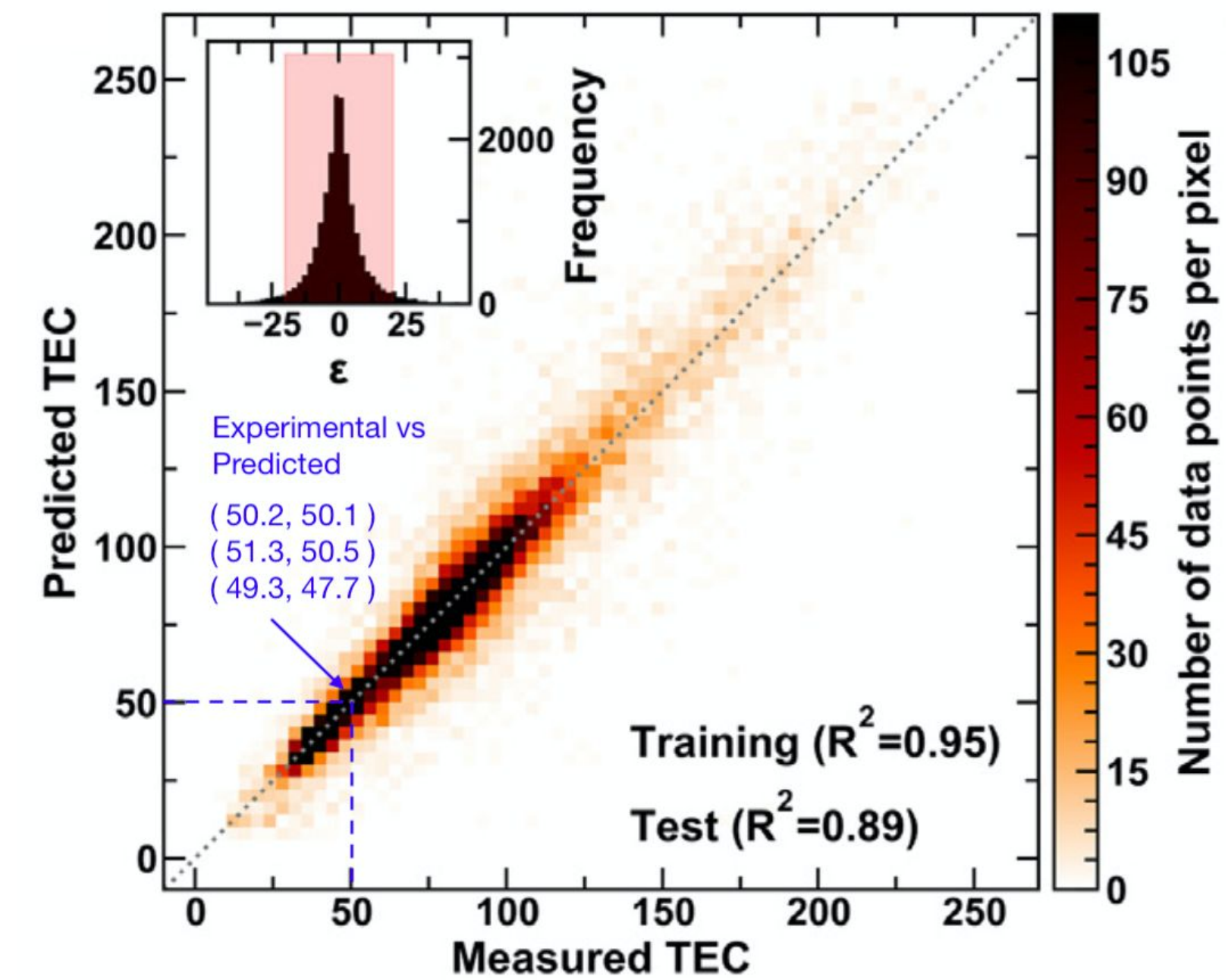
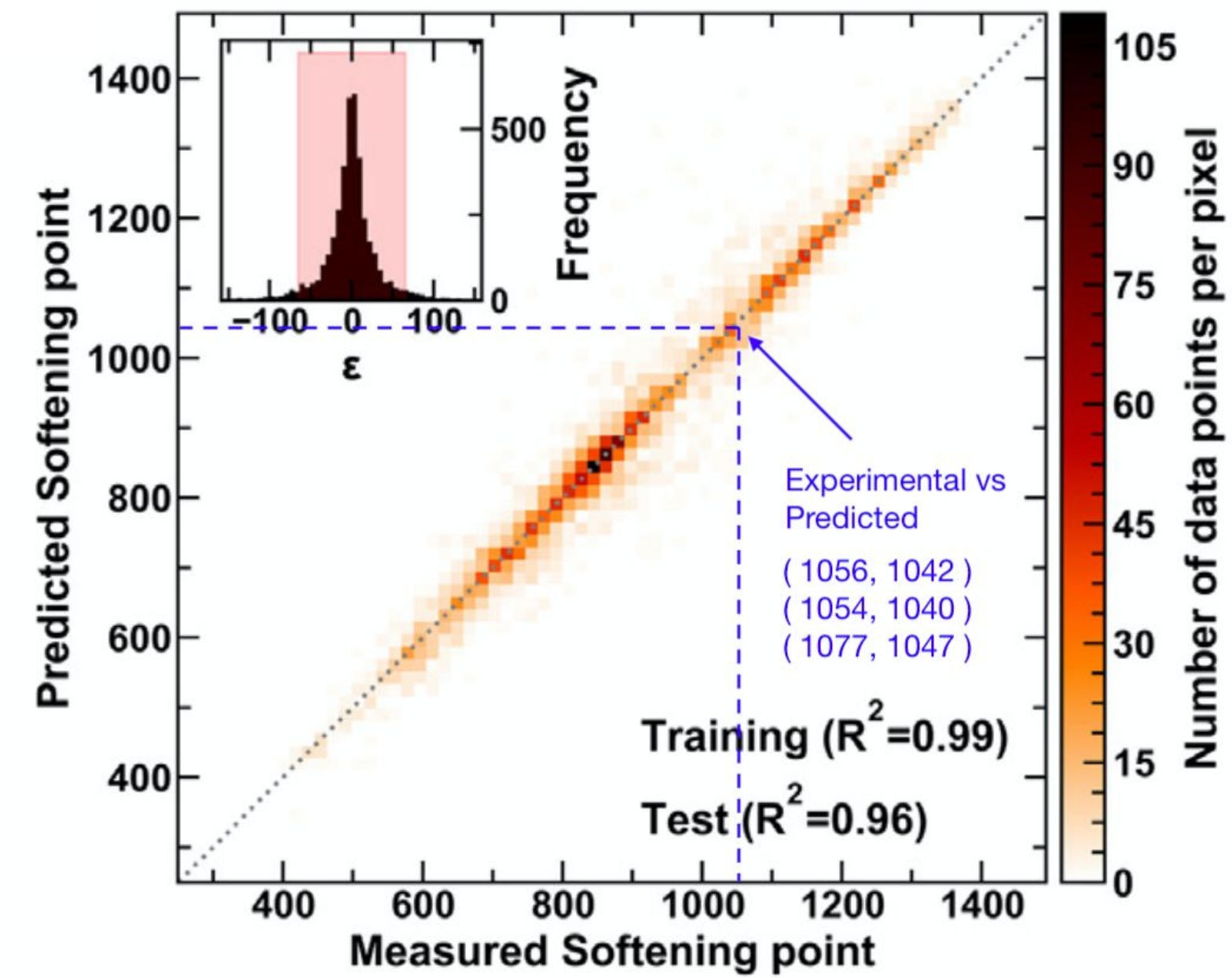
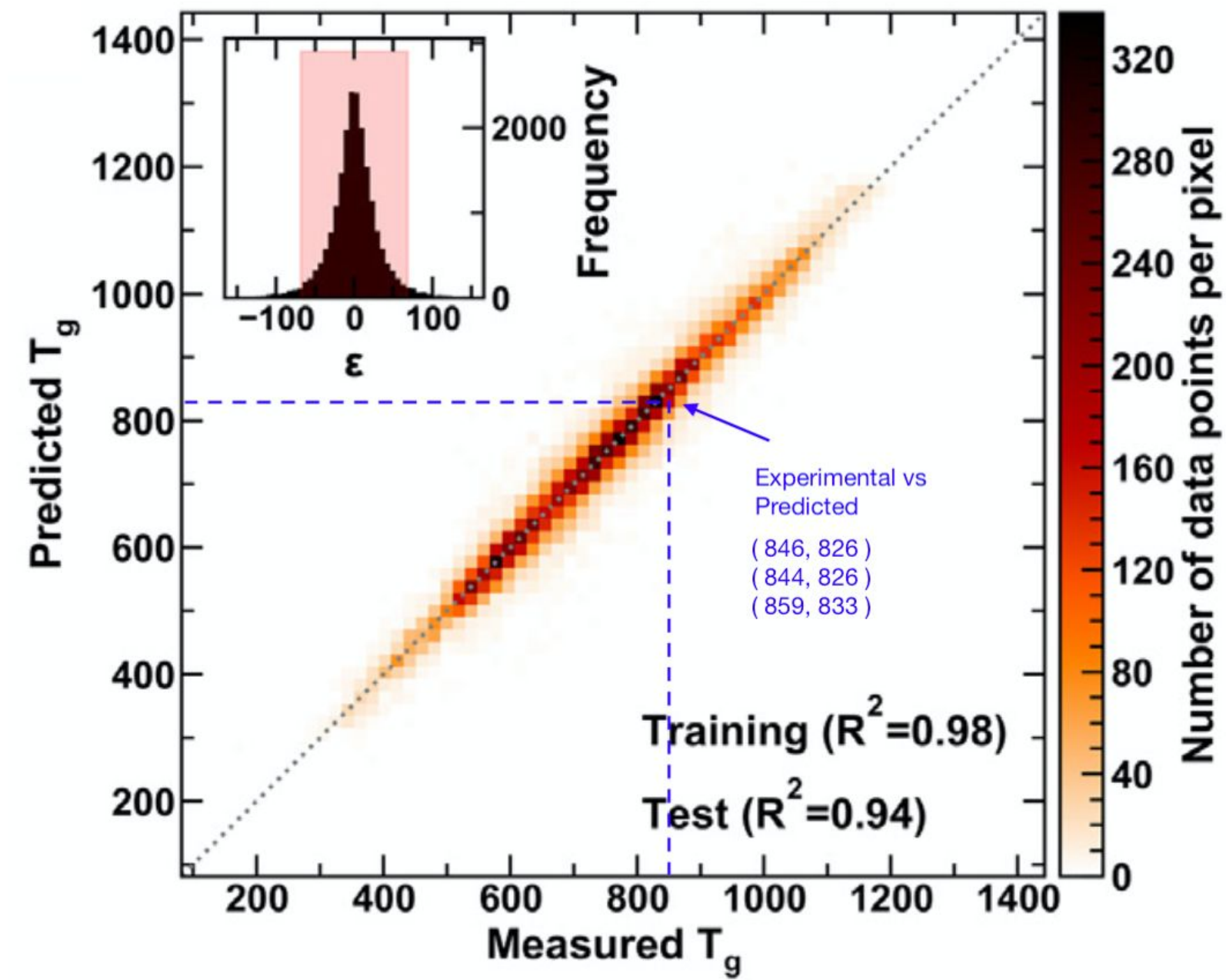
Property Constraints

Property	Range
Glass Transition Temp.	833 – 843 K
Coeff. of Thermal Expansion	52 – 58 (x10 ⁻⁶)

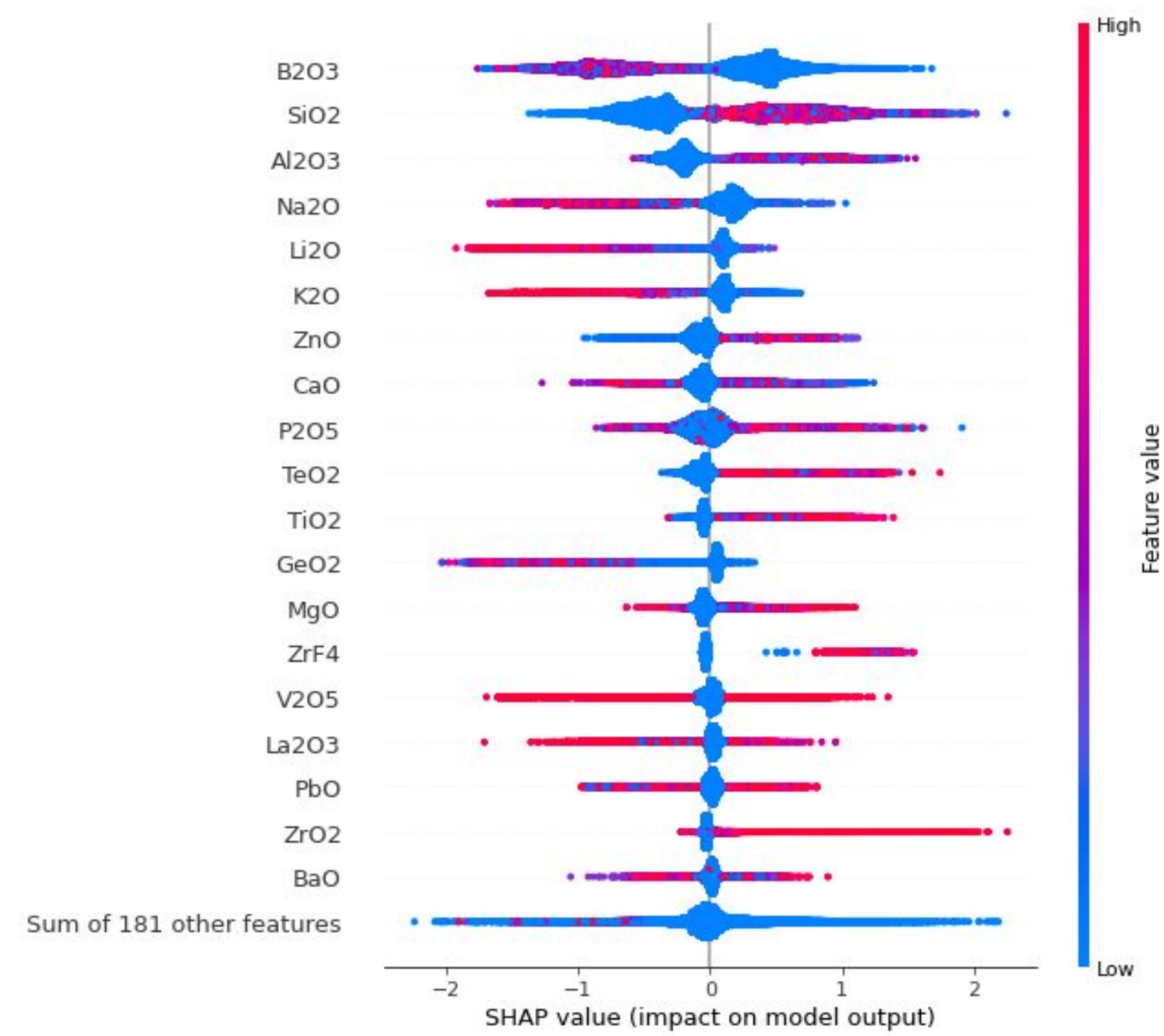
Compositional constraints

Component (wt%)	Range
SiO ₂	67 - 77
B ₂ O ₃	5 - 15
Al ₂ O ₃	1 - 9
CaO	0 - 4
Li ₂ O	0 - 1
Na ₂ O	1 - 13

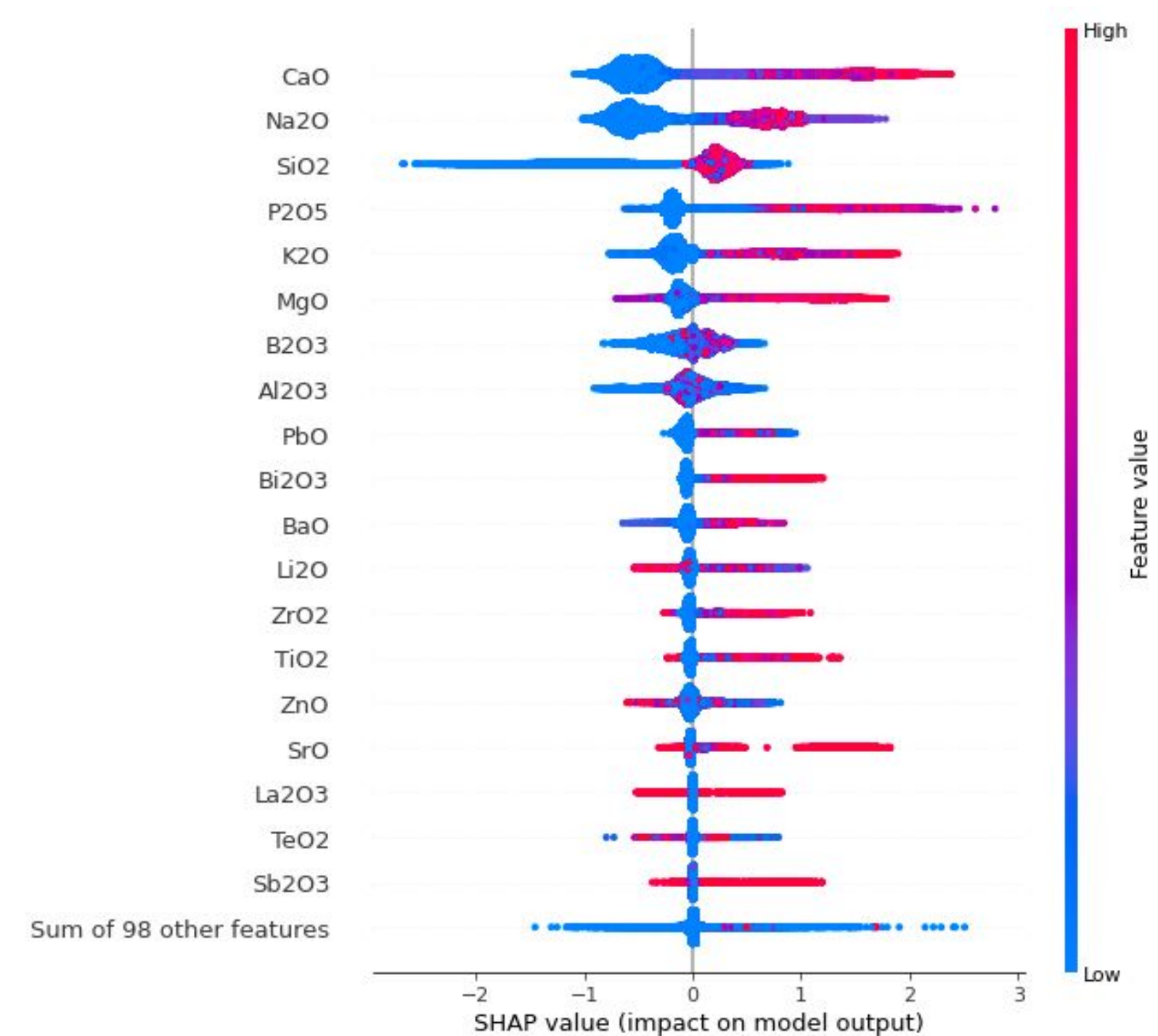
Constraints for pharmaceutical glass (Sisecam)



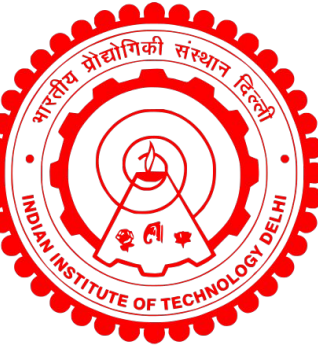
Interpretability



T_g



TEC



Towards Industrial Scale

Automation at Plant Level?

Cement manufacturing

- Cement manufacturing: an extremely complex process
- ~8% global CO₂ emissions
- Develop a digital twin for process optimization and quality control in collaboration 5 cement companies round the globe

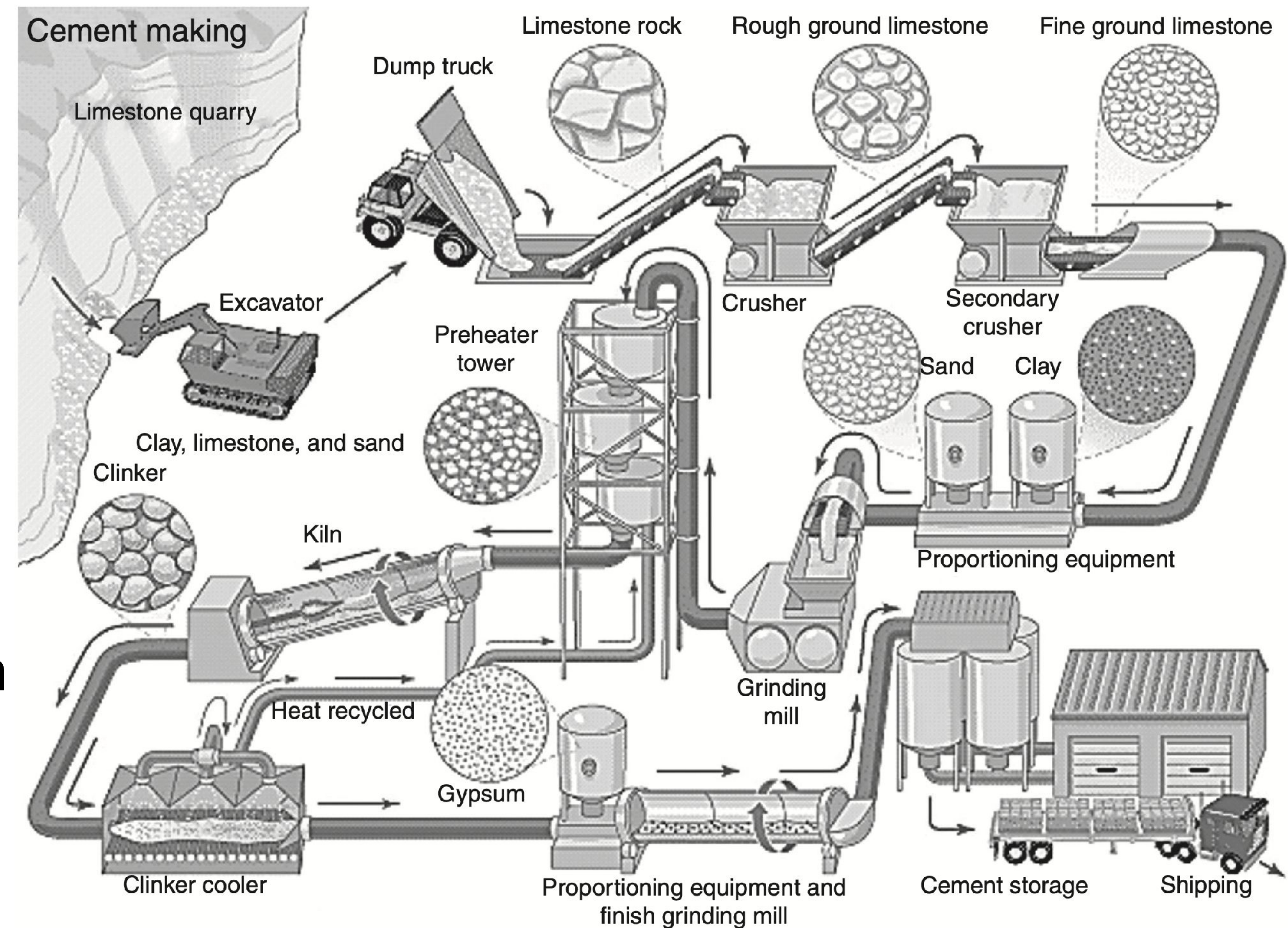
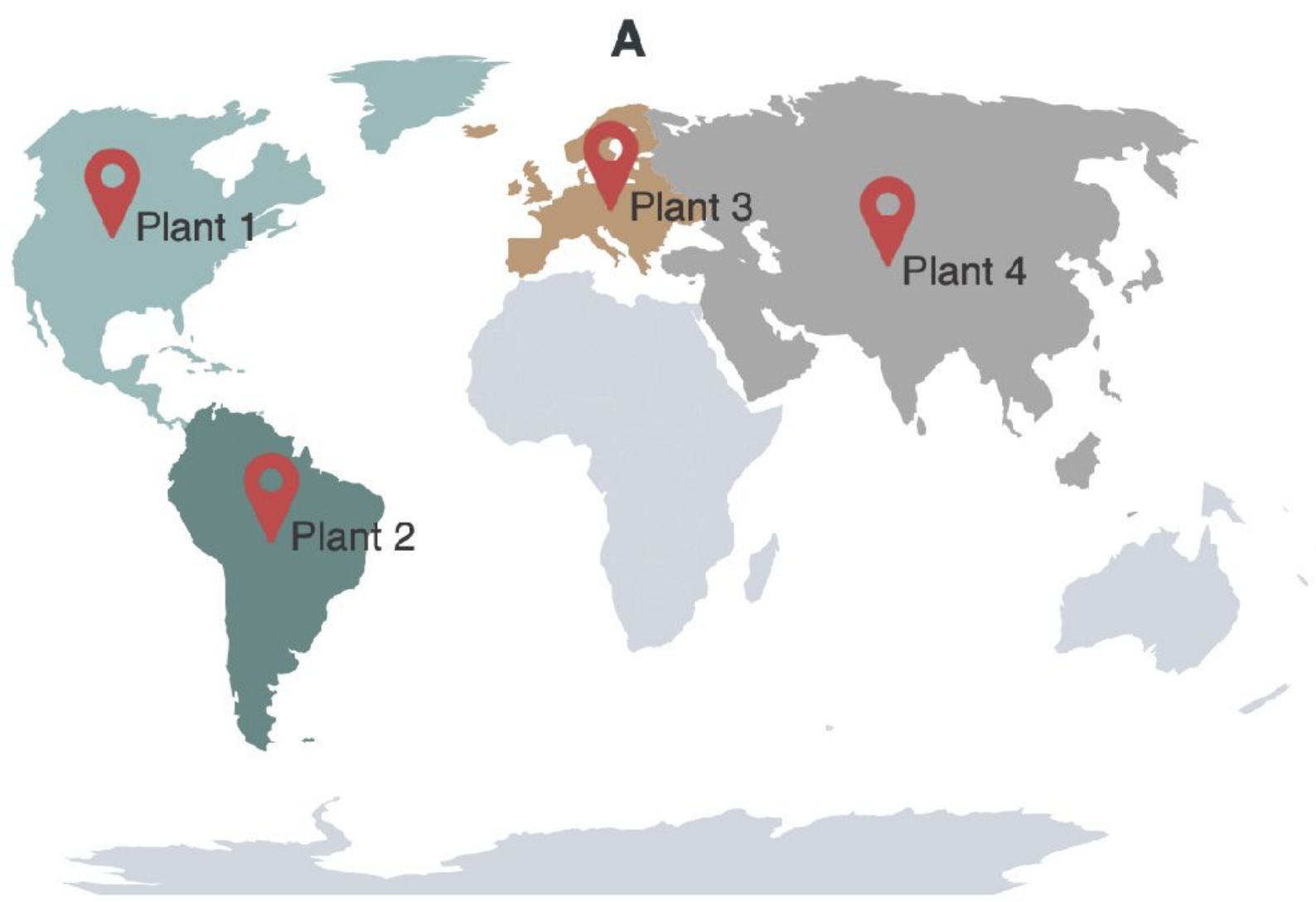


Image Source: Encyclopædia Britannica, Inc.

Data curation

Diversity : geography, fuel mix, configuration

Different levels of **data richness** :
Temporal granularity & data volume



Cement plant specifications				Data specifications							
ID	Location	Configuration	Fuel mix	Emissions measured (PPM)	Measurement location	Sampling frequency	Data Duration	Size ‡			
								Raw	After consistency check	After physical validation	After outlier removal
1	North America	5 stage PHT + ILC pre-calciner	Coal	NO _x CO CO ₂	kiln inlet† (NO _x) stack end (CO) stack end (CO ₂)	1 minute	01 Jan 20 - 31 Dec 21 (2 year)	1,052,567	894,036	277,603	277,079
2	South America	5 stage PHT + ILC pre-calciner	Petcoke, charcoal, waste wood chip, corn husk, rice husk, residual tires	NO _x	kiln inlet	2 hour	11 Jan 20 - 31 Dec 21 (2 year)	6,907	6,907	6,907	5,994
3	Europe	Reinforce suspension preheater (RSP) tower	Petcode, RDF	NO _x	kiln inlet	1 hour	30 Nov 21 - 22 Dec 22 (1 year)	5,784	2,621	2,621	1,620
4	Asia	6 stage PHT + ILC pre-calcienr	Petcoke, coal	CO	Stack end	1 hour	01 Feb 22 - 01 May 22 (3 months)	2,165	1,389	1,217	1,151

Incorporating kiln-dynamics

$$y(t) = f_{EP}[PP(t)]$$

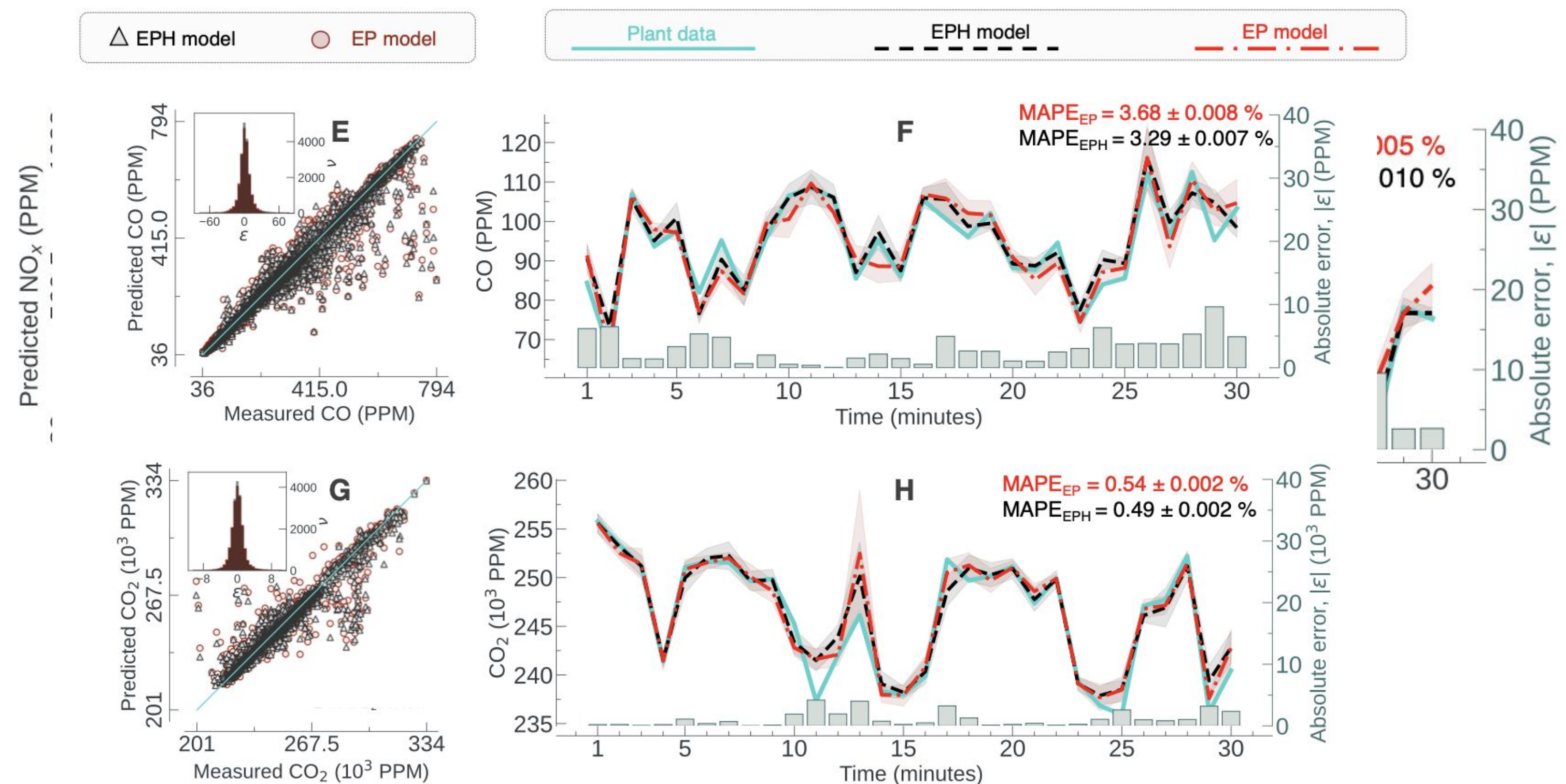
$$y(t) = f_{EPH}[PP(t), PP(t-1), \dots, PP(t-\tau)]$$

τ : Process history

Incorporate 20 min of process history
(τ)

~3x reduction in NO_x error :
5.7% (EP) vs. 1.99% (EPH)

Dependence on prior operation



NO_x Control framework

DVs: Controllable process parameters

Constraints:

$$\Delta |DV| \leq 5\%$$

$$\Delta RM_{\text{flow chemistry}} = 0$$

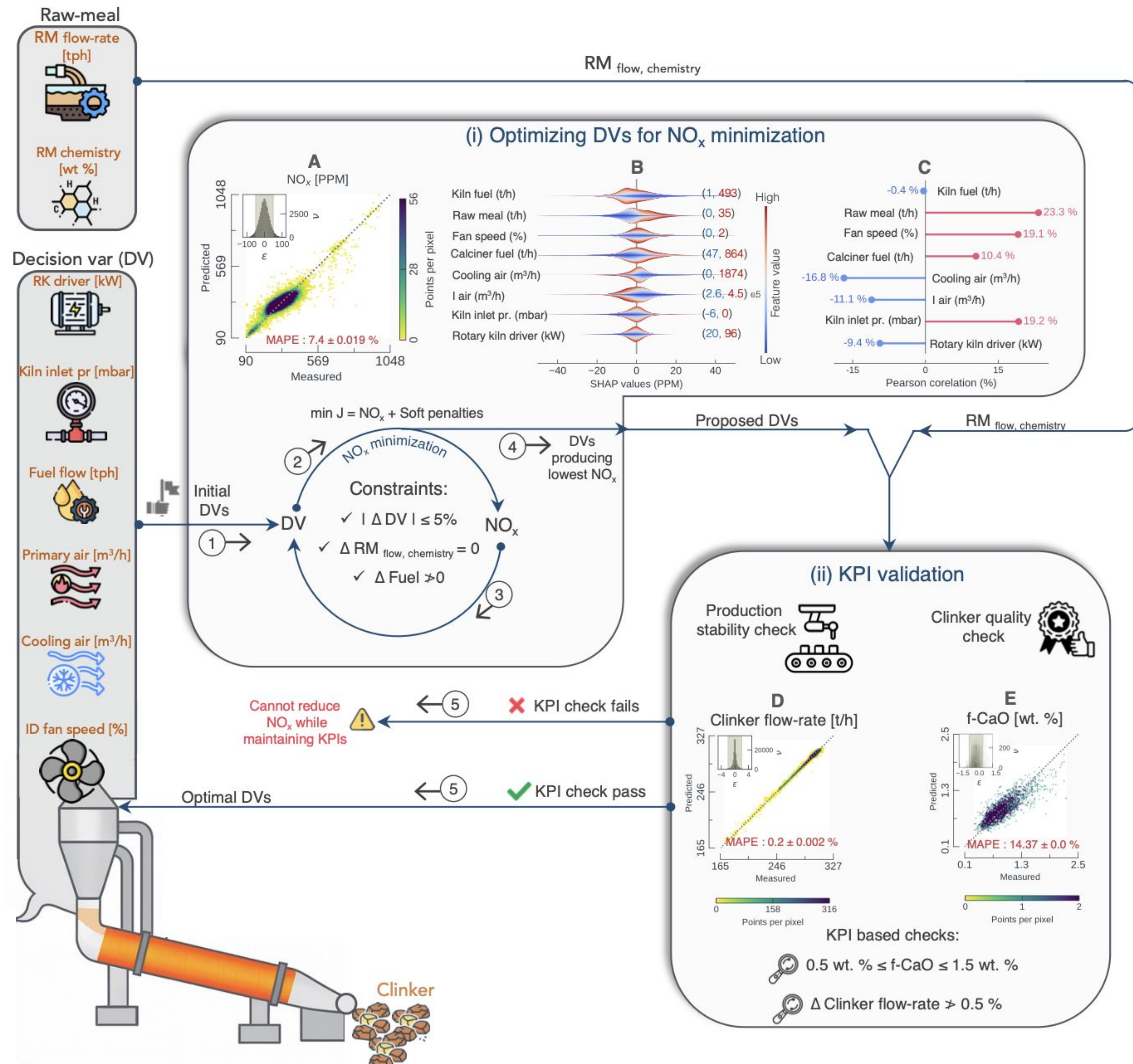
$$\Delta \text{Fuel} \leq 0$$

Objective function penalties:
Consistent with correlation & operational feasibility

KPI validation:

$$f\text{-CaO} \in [0.5 \text{ wt. \%} - 1.5 \text{ wt.\%}]$$

$$\Delta \text{Clinker flow} \leq 0.5\%$$

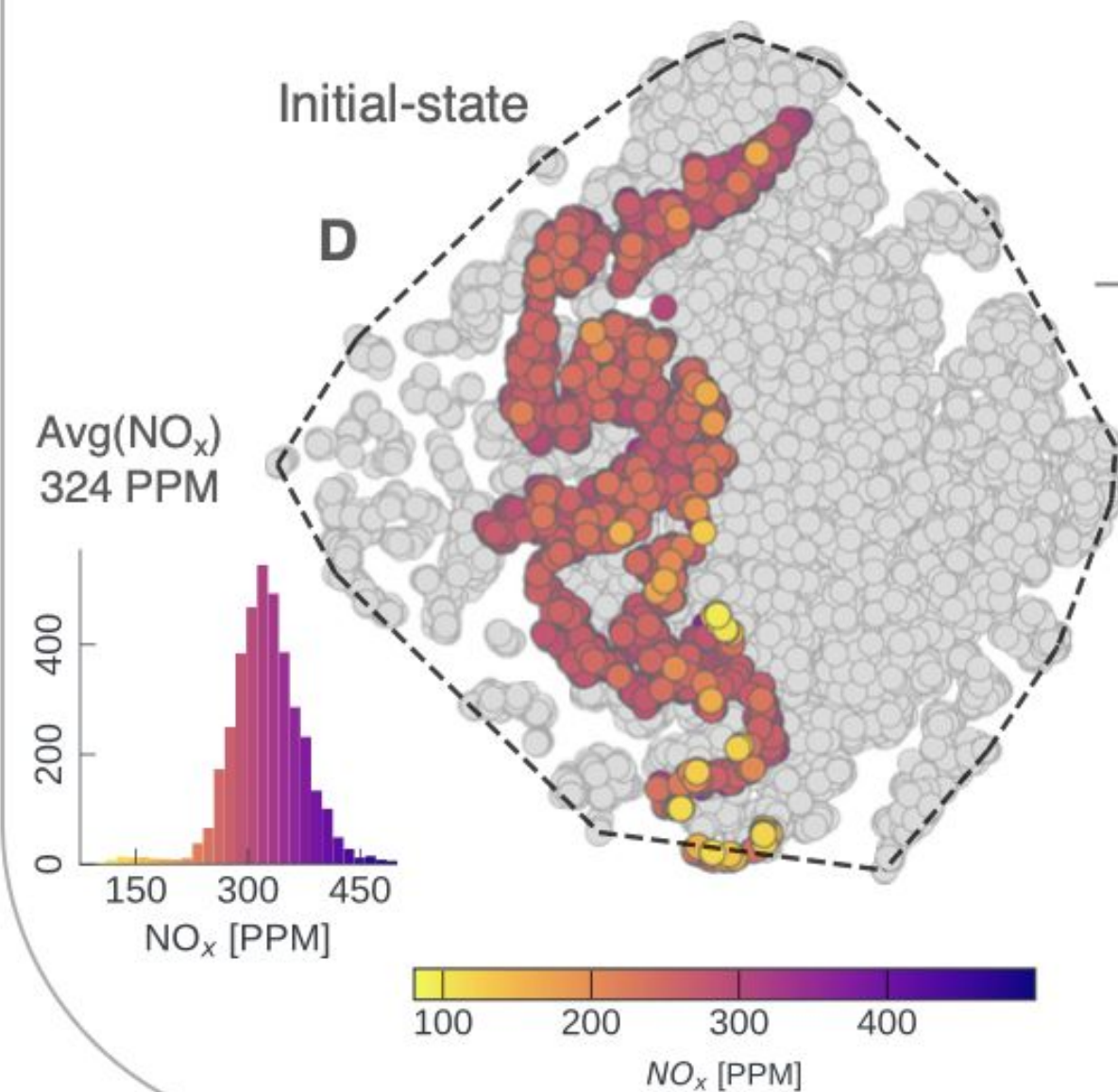
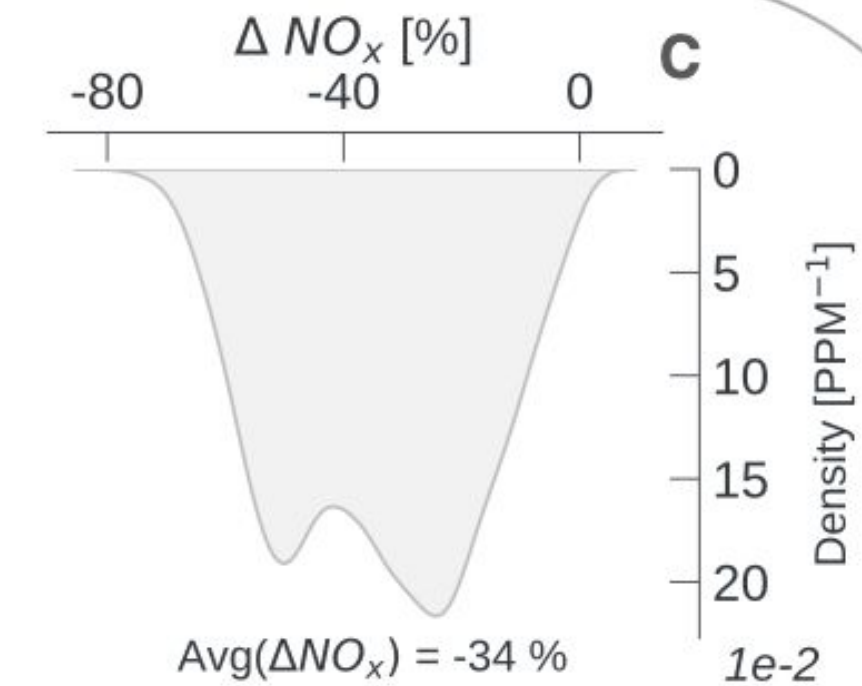
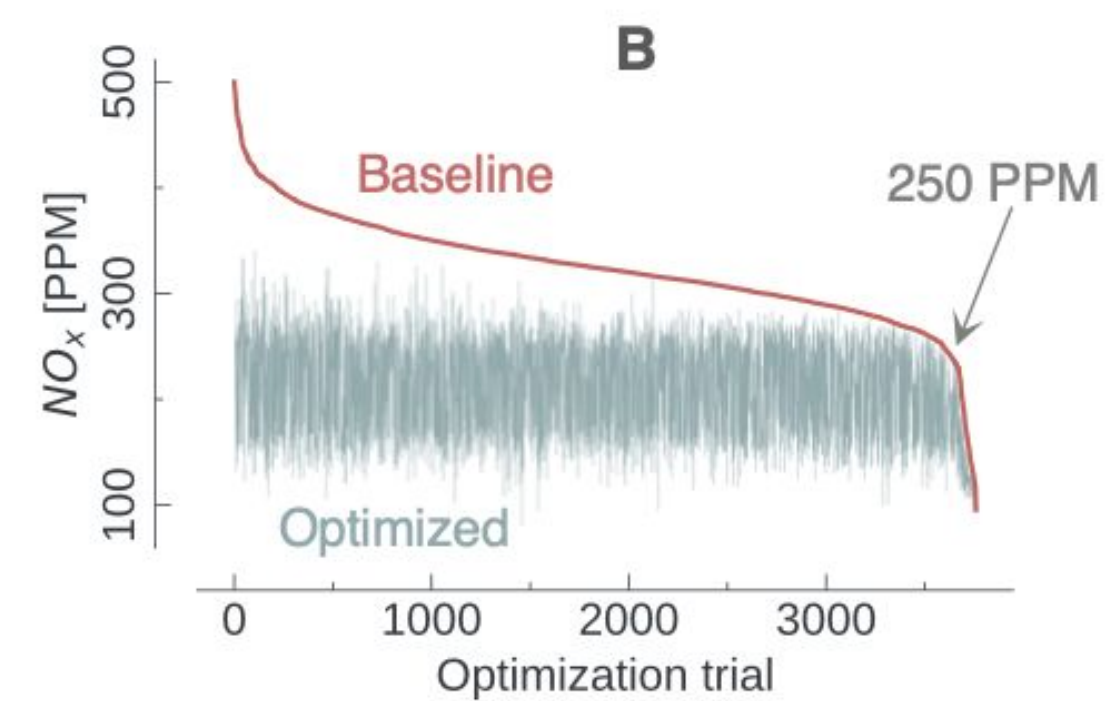
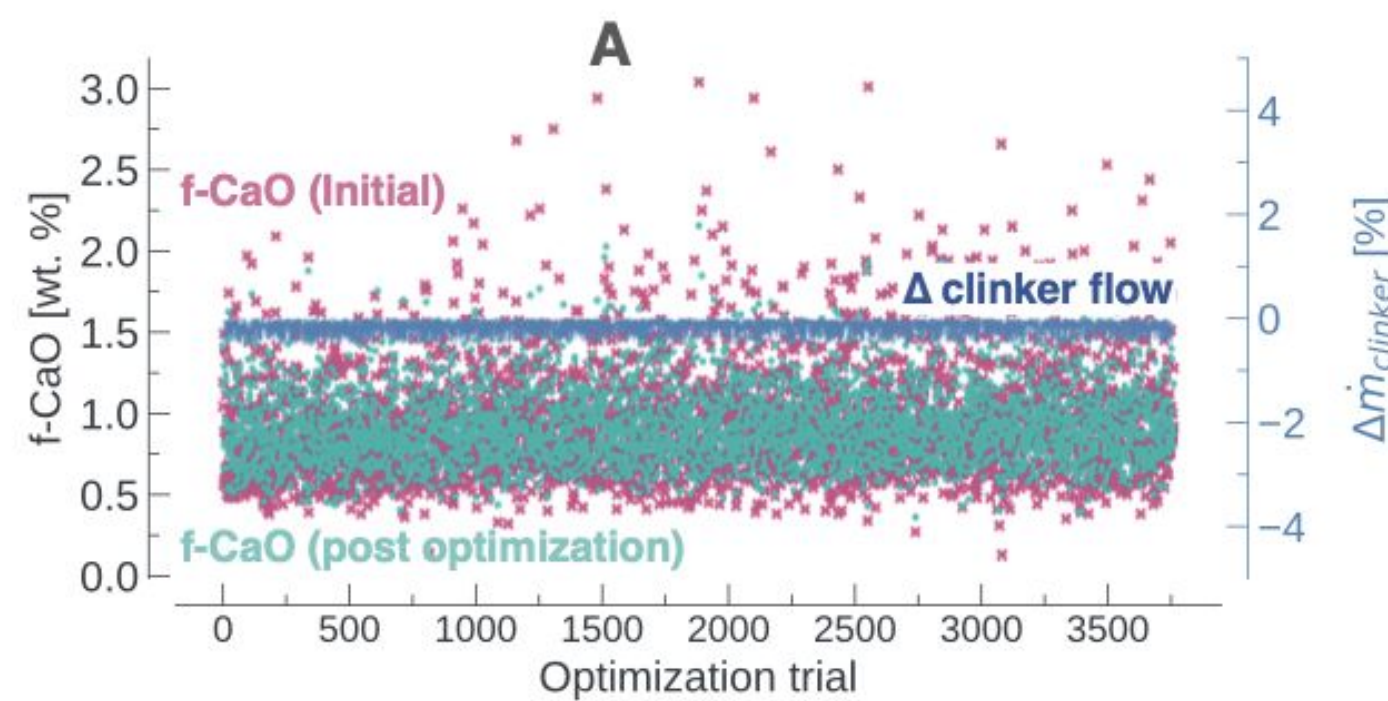


Controller performance

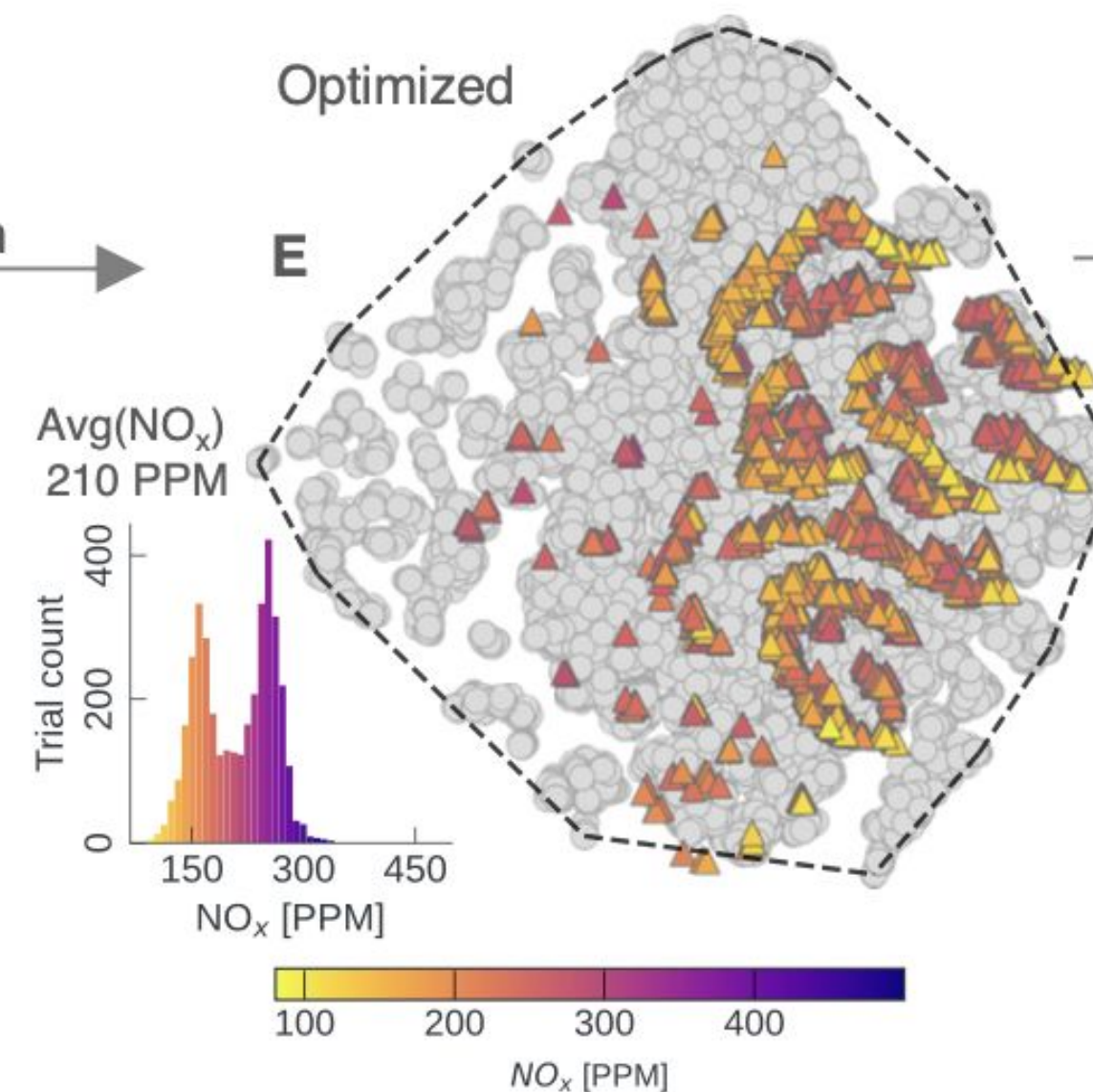
34% reduction

Scenario 1: Normal plant operation ($\text{NO}_x \leq 500$ ppm)

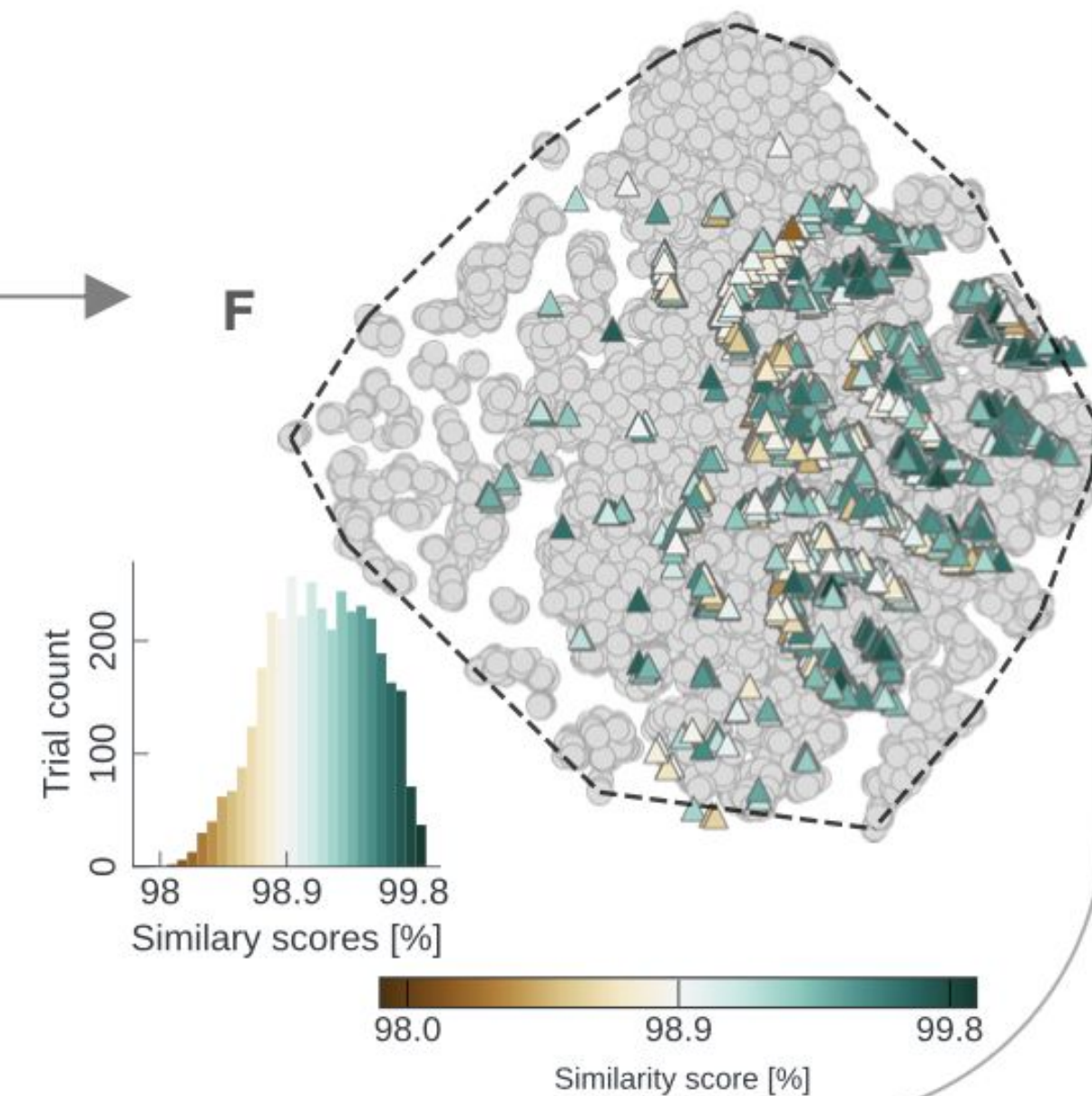
Normal plant
Operation
(3,761 trials)



Optimization



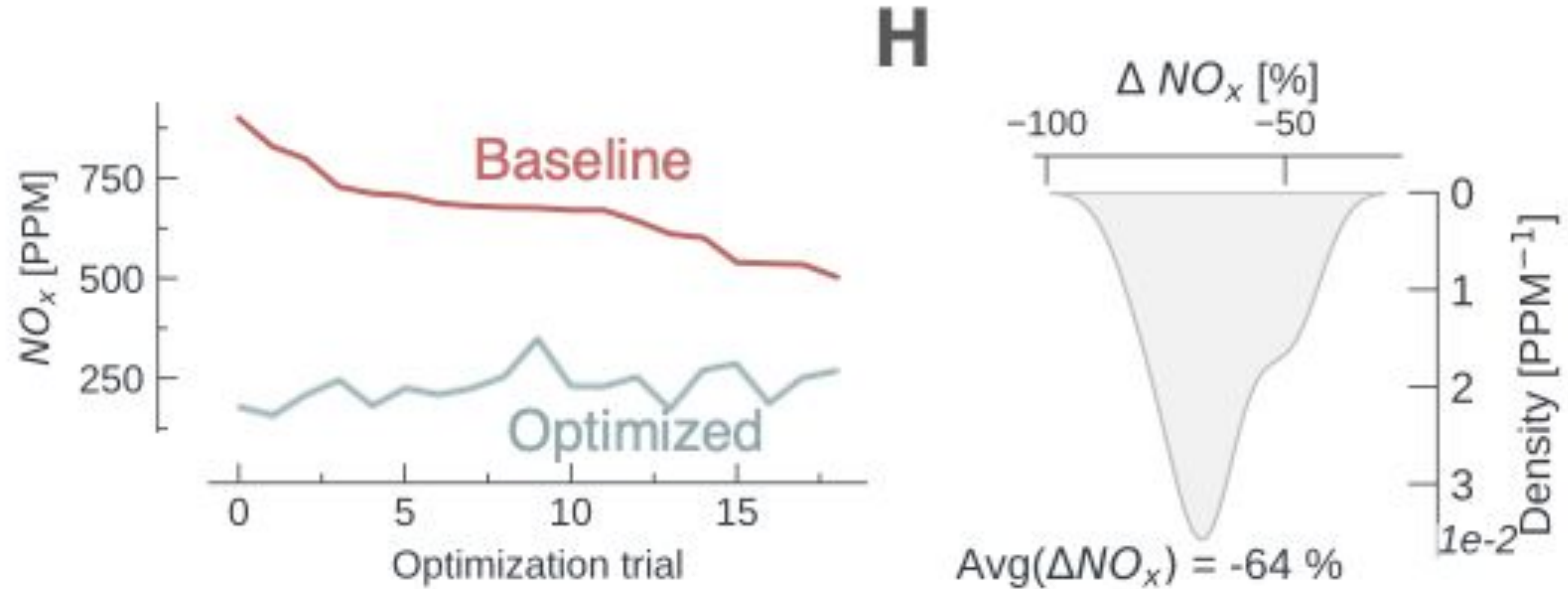
Validation



Controller performance

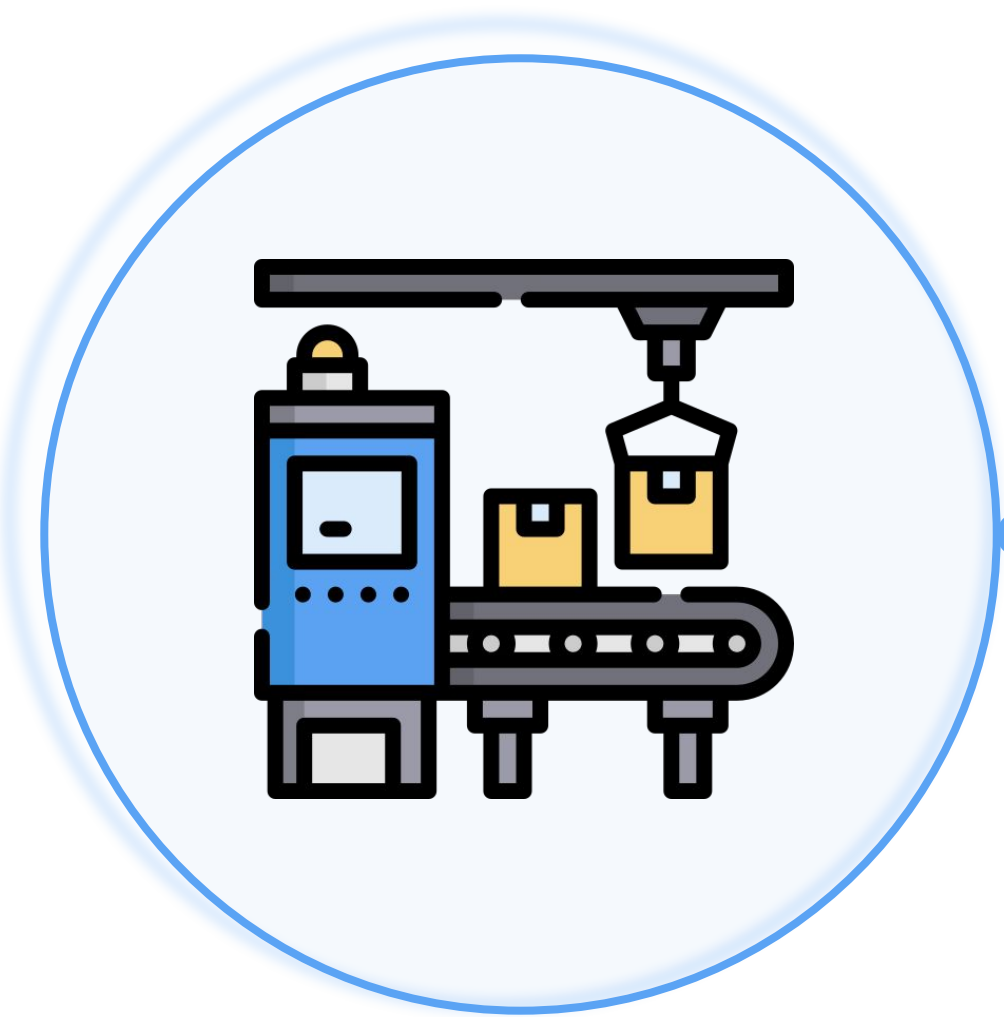
64%
reduction

Scenario 2: Stress test ($\text{NO}_x > 500$ ppm)



Techno-economic evaluation of the controller

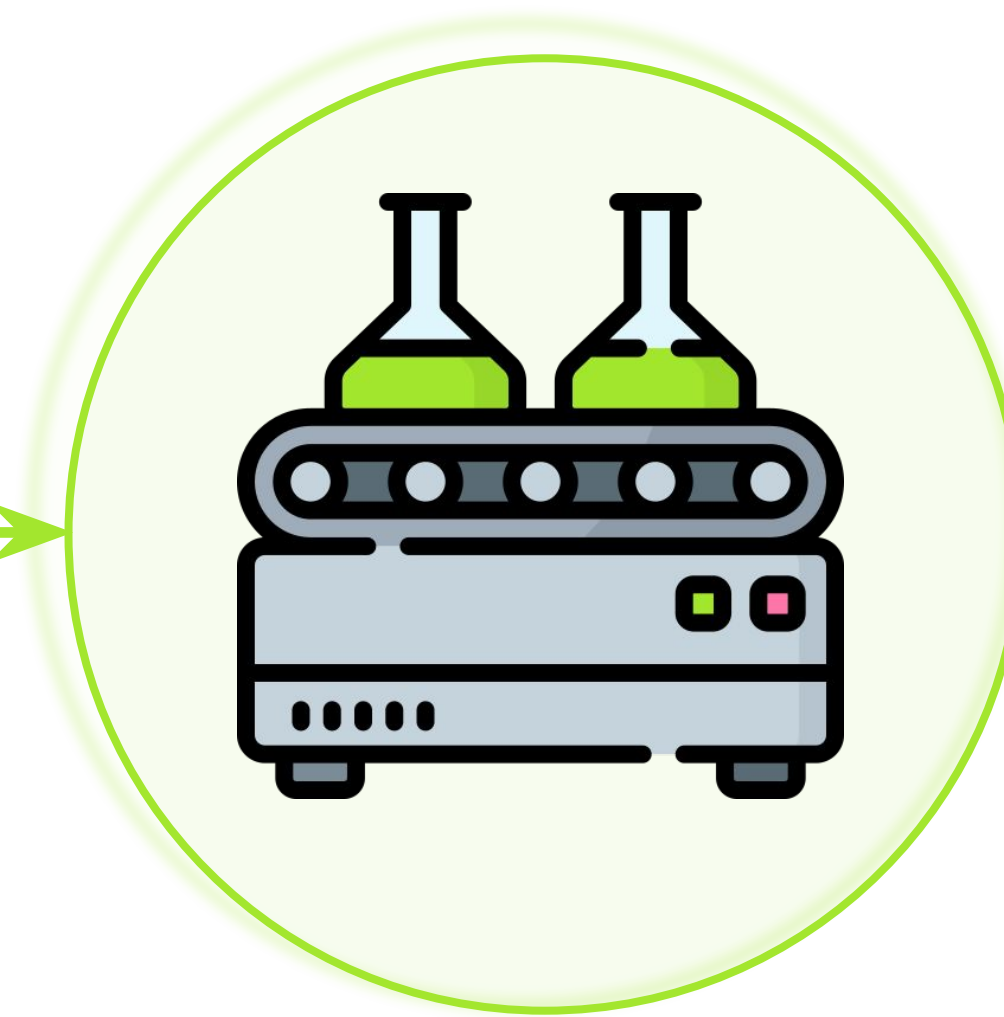
2.3 Mt clinker/yr
production



~290 t NO_x/yr
controlled



~129 t NH₃/yr
avoided



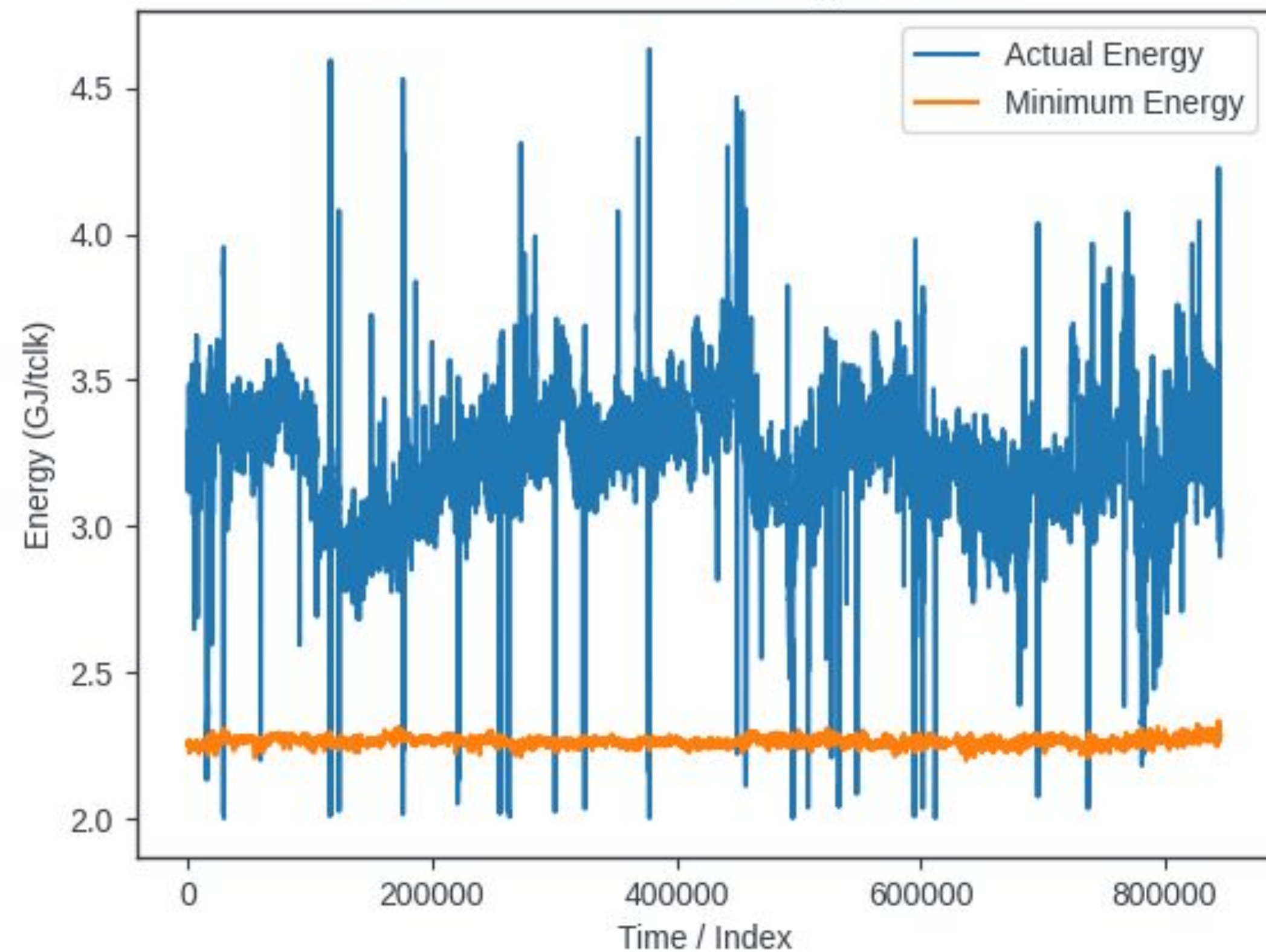
~58,000 USD/yr
saved



[1] U.S. Environmental Protection Agency, Selective Noncatalytic Reduction (SNCR) for NO_x Control: Technical and Cost Guidance, U.S. EPA, chapter 1 (SNCR process description, stoichiometry, NSR and reagent calculations) (2017).

Extension to fuel & CO₂ minimization

Actual and Minimum Energy Over Time

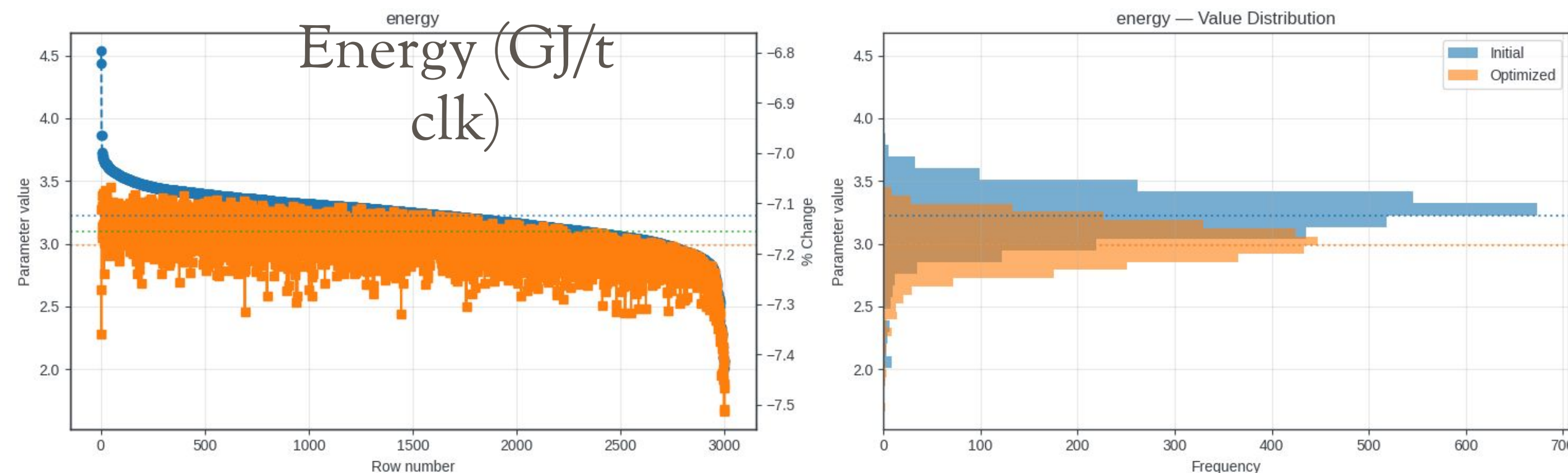


Mean actual = 3.2 GJ/ton clinker

Mean theoretical minimum = 2.3 GJ/ton clinker

Fuel reduction = 7.15%

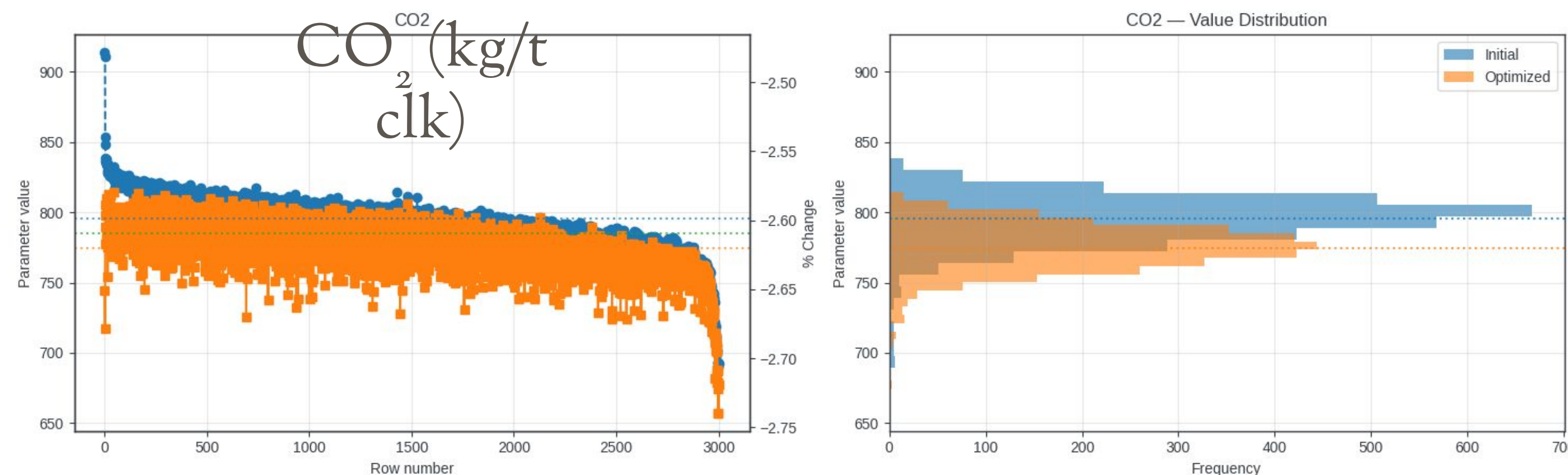
Mean Initial: 3.226 | Mean Optimized: 2.992 | Mean % Change: -7.15%



Corresponding CO₂ reduction =

2.61%

Mean Initial: 795.277 | Mean Optimized: 774.413 | Mean % Change: -2.61%



Agentic Decision Support System

STATUS: Dashboard live! Last data received: 2022-09-12 22:00:21

Kiln Feed

252.4

tonnes

MB Fuel

10.88

tonnes per hour

PC Fuel

15.9

tonnes per hour

Fan Speed

75.5

percentage %

Preheater Outlet O₂

4.2

percentage %

SHC

834.5

kcal per kg

Blockage Risk

Normal

mbor

RECOMMENDED ACTION

Coal Main Burner

NA

tonnes per hr

Accept action

RECOMMENDED ACTION

Coal Calciner

NA

tonnes per hour

Accept action

RECOMMENDED ACTION

Kiln Feed

NA

tonnes/hr

Accept action

SHC

NA

kcal per kg

Predicted Clinker C3S

NA

percentage

Predicted Clinker FCaO

NA

percentage

Operator Feedback

Send Feedback

Plotly Cloud

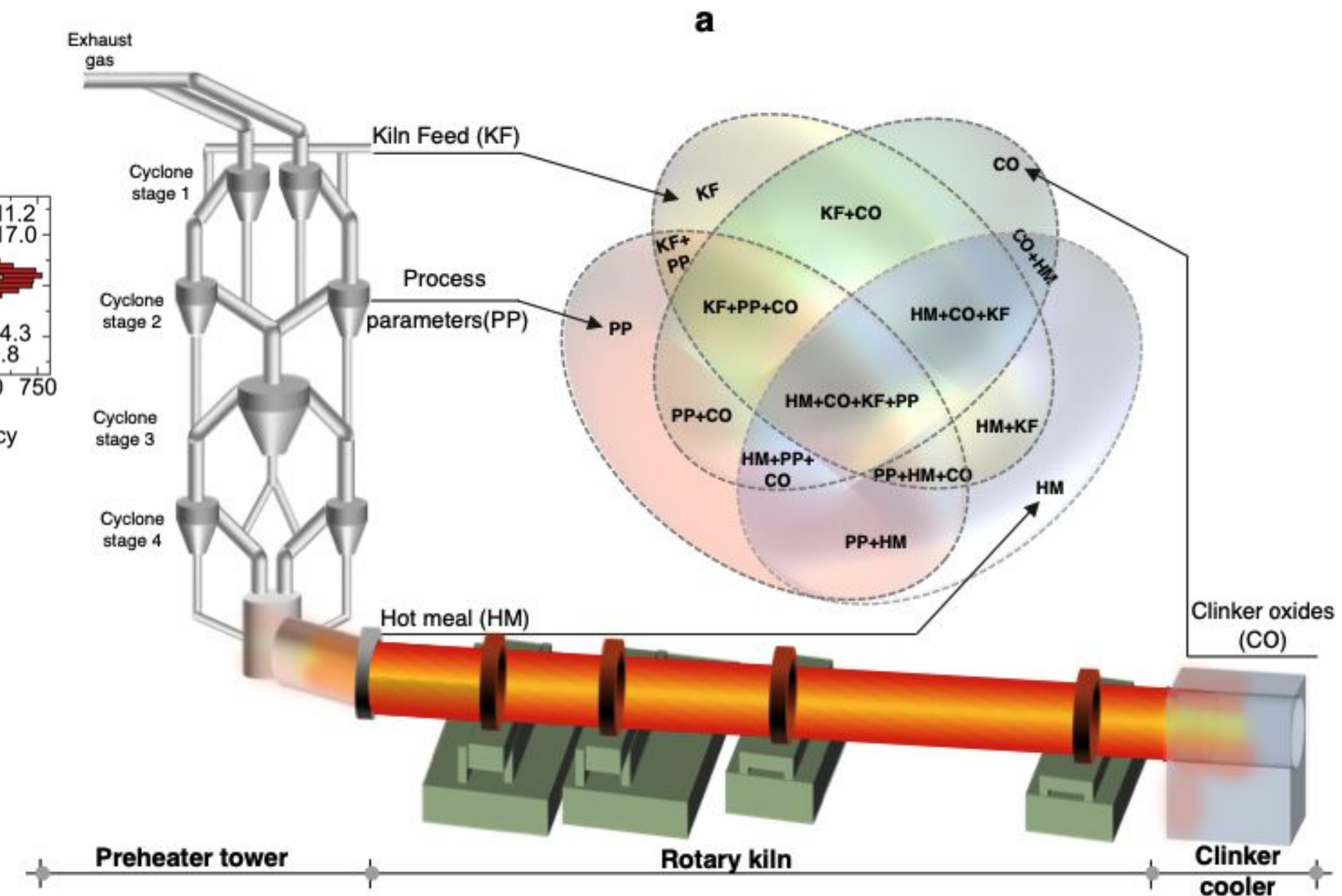
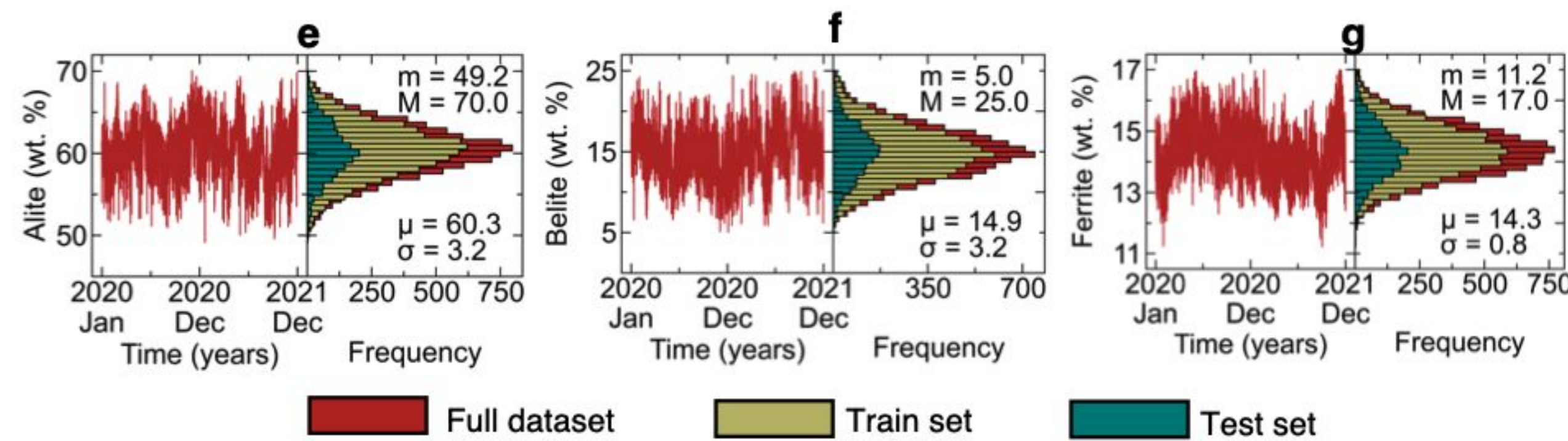
Errors

Callbacks

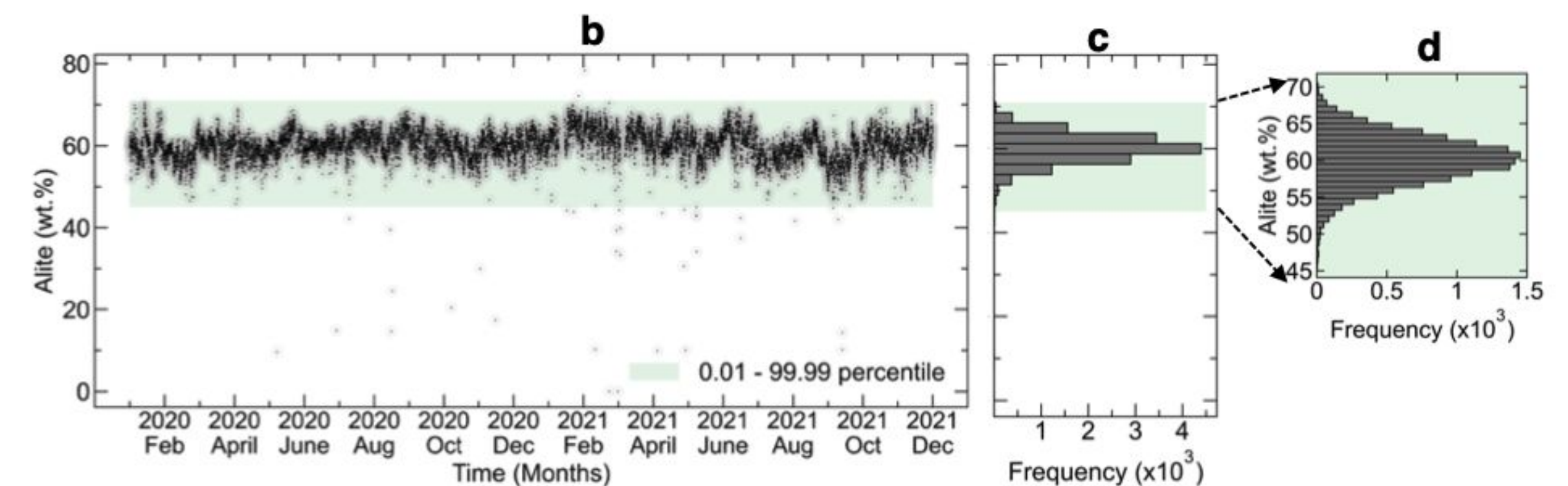
v3.4.0

Server

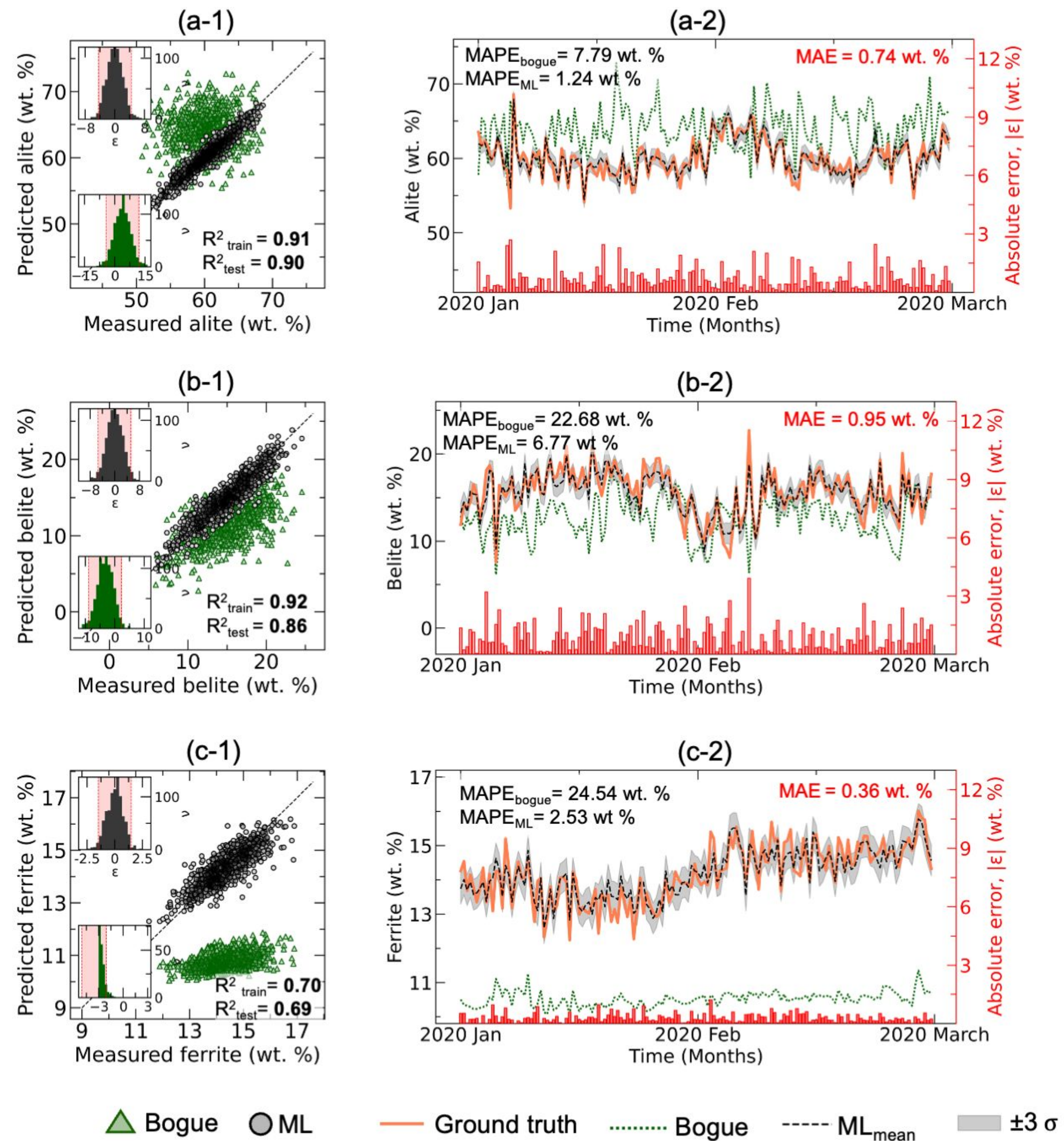
Model features



- Input features include raw materials, fuel, process parameters
- Goal: Clinker mineralogy
- Major challenge: Data cleaning and alignment



Model performance



- Accurately forecast clinker compositions
- Allows direct control of the plant for optimization
- Performs better than the models employed in the plant

Way forward: Proposal

Industry Academia joint Center of Excellence/Consortia on Glass

- Goal: Towards Sustainable Glass Manufacturing and leadership in Glass Industry
- Problem statements of industrial relevance proposed by Industry
- Mentorship from both academia and industry
- What we have: Strong industry, world-class research, increasing demand (unlike several other parts of the world)



Mausam



K. Jablonka



Indrajeet Mandal



Vinay Kumar



Sajid Mannan

Thank you!



intel
labs



High Performance Computing at IITD

Supercomputer @ IITD : The Central Hybrid Supercomputing Cluster at Indian Institute of Technology Delhi



Alexander von
HUMBOLDT
STIFTUNG



Google DeepMind



Undergraduate and doctoral researchers at M3RG, IIT Delhi

Thank you!



High Performance Computing at IITD

Supercomputer @ IITD : The Central Hybrid Supercomputing Cluster at Indian Institute of Technology Delhi

