



25.01.2024

DATA DRIVEN MATERIALS DESIGN PART III.

Dr. Christoph Völker

christoph.voelker@bam.de



Scope



- Introduction to data driven materials design (last week)
 - Digital materials
 - Predicting properties
 - Decision making
- 2. Prospects of LLMs
- 3. Hands on (next week)
 - AI-driven materials design
 - SLAMD
 - LLM

Ressources



Materials Design App

SLAMD: https://slamd-demo.herokuapp.com/



Data: https://github.com/BAMcvoelker/Praktikum_MD/

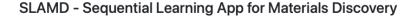
- SessionExample.json <- a SLAMD session file that contains some of the materials
- SessionComplete.json <- a SLAMD session file that contains the complete list of materials and materials blends that are required to create the geopolymer based concrete formulation
- DiscoveryData_Sample.csv <- a materials data search space with labels
- Slides





SLAMD Workflow

www.bam.de





Leverage the Digital Lab and AI optimization to discover exciting new concrete recipes

- > Represent resources and processes and their socio-economic impact.
- > Calculate complex compositions and enrich them with detailed material knowledge.
- > Integrate laboratory data and apply it to novel formulations.
- > Tailor concretes to the purpose to achieve the best solution.

Workflow

Digital Lab



1. Specify resources

From base materials to manufacturing processes - "Base" enables a detailed and consistent description of existing resources.



2. Combine resources

The combination of base materials and processes offers an almost infinite optimization potential. "Blending" makes it easier to design complex configurations.



3. Digital recipes

With "Formulations" you can effortlessly convert your resources into the entire spectrum of possible concrete formulations. This automatically generates a detailed set of data for AI optimization.

2. AI-Optimization



4. Materials Discovery

Integrate data from the "Digital Lab" or upload your own material data. Enrich the data with lab results and adopt the knowledge to new recipes via artificial intelligence. Leverage socio-economic metrics to identify recipes tailored to your requirements.

Link:

it Powered by iteratec

https://slamd-demo.herokuapp.com

Links







Digital Lab



Digital Twin



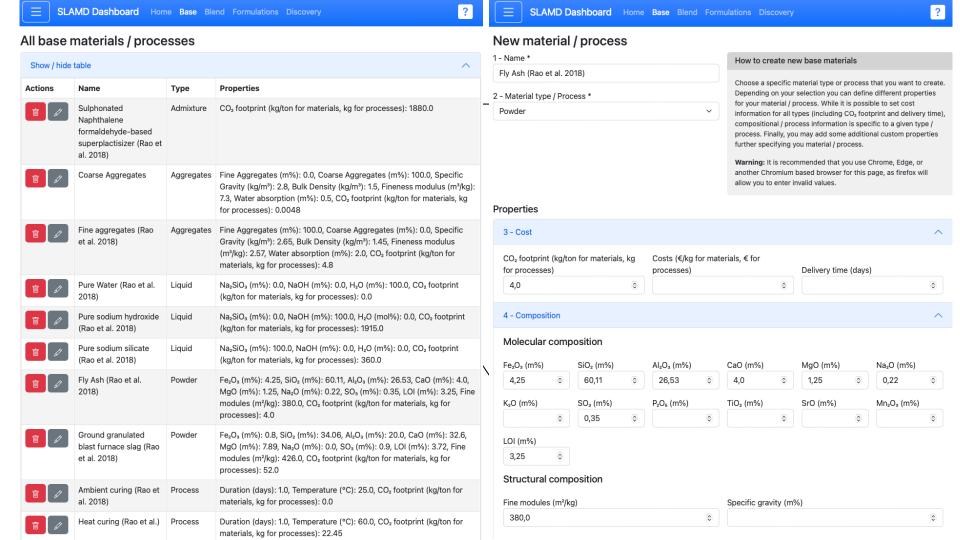


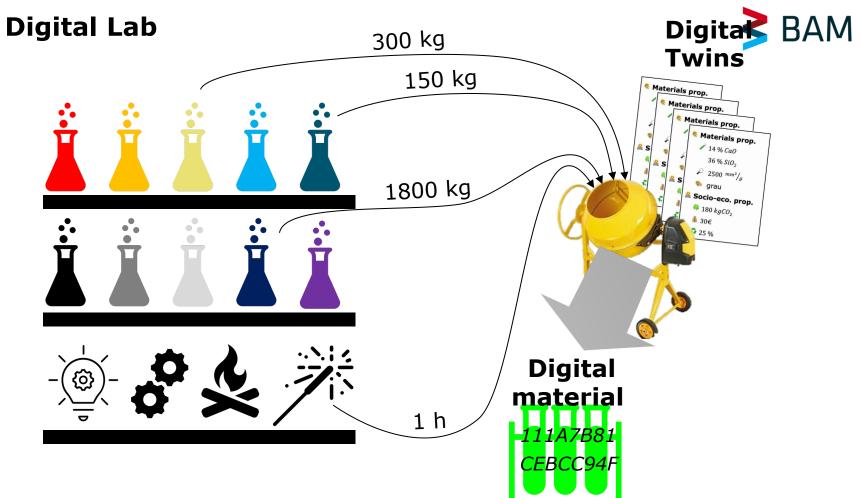


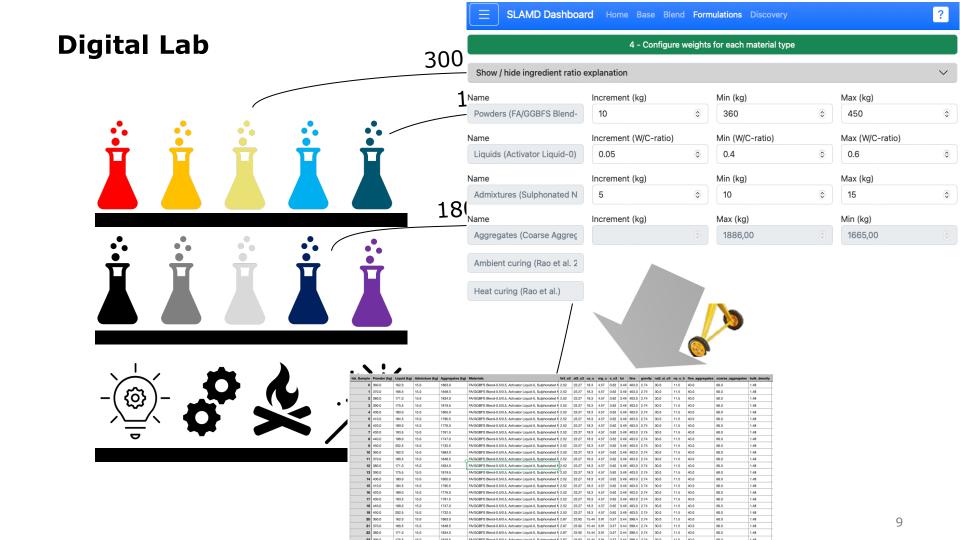
- Materials prop.
 - *✓*% CaO

..... % SiO_2

- $p = mm^2/g$
- Socio-eco. prop.
 - ***** *kgCO*₂
 - <u>\$</u>€
 - **\$** %



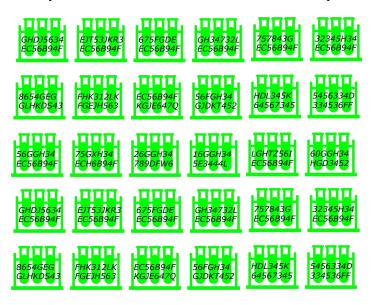




AI Optimization - Materials discovery dashboard



Import materials search space



Set objectives



Strenght (Maximize)
Threshold: 45 MPa

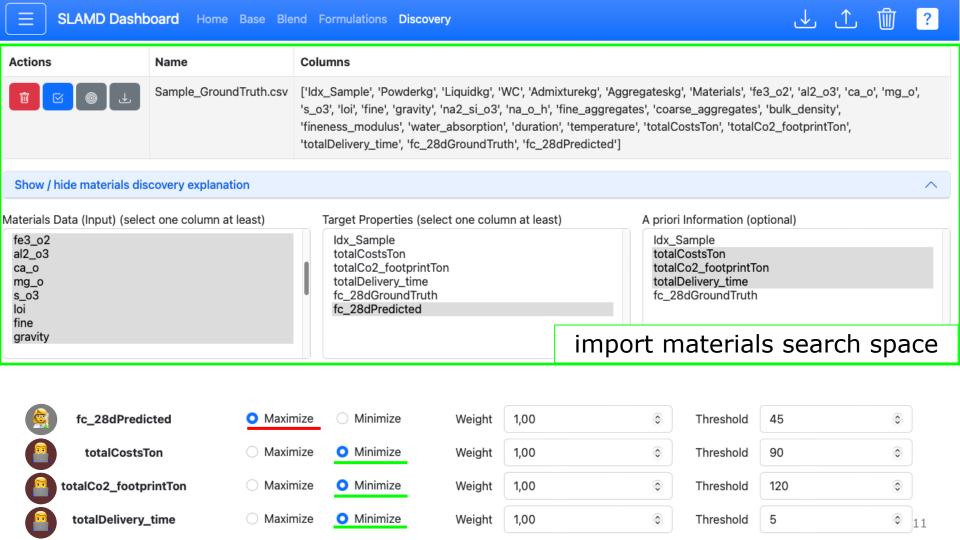


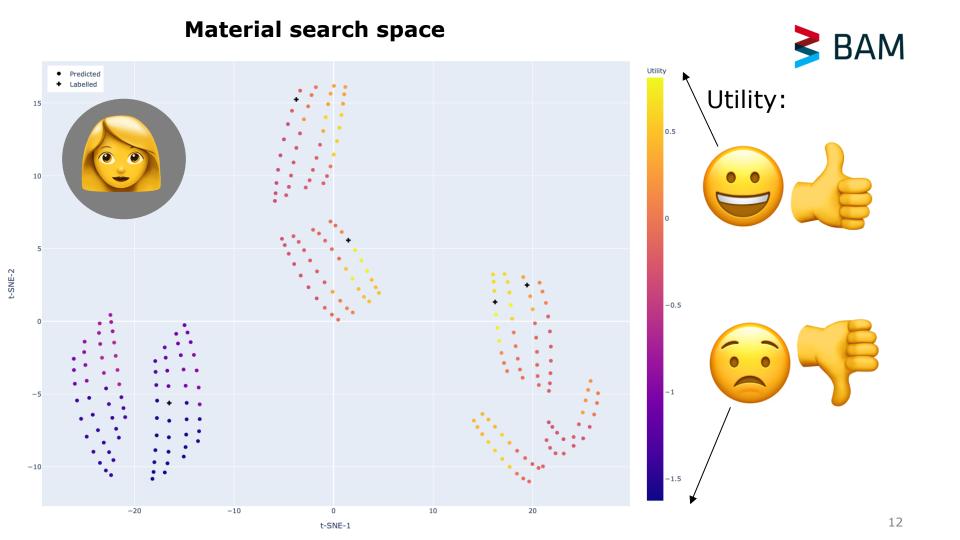
Costs (Minimize)

Threshold: 90 €

CO2-Footprint (Minimize)

Threshold: 120 kg/ton







How to

www.bam.de

How to?



The configuration of the optimization problem is the key to its success. Here is some advice base on a large-scale literature study where we have probed questions such as:

Q1: Which **ML-algorithm**?

Q2: How much **training data**?

Q3: Does **informed AI** improve performance?

9 search spaces*

1... 10...20 ref's from literature

90 ... 152 ... 274 recipes

28 ... 37 ... 43 parameters



compressive strength (MPa)



CO₂ footprint (kg/ton)

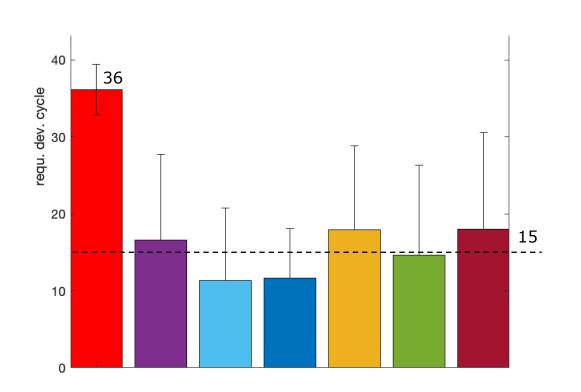


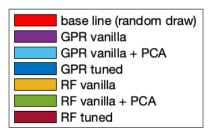


Benchmarking (required development cycle)

Q1: Algorithm







Mean required development cycle to reach top 5% performance with 90% certainty.

9 data sets *

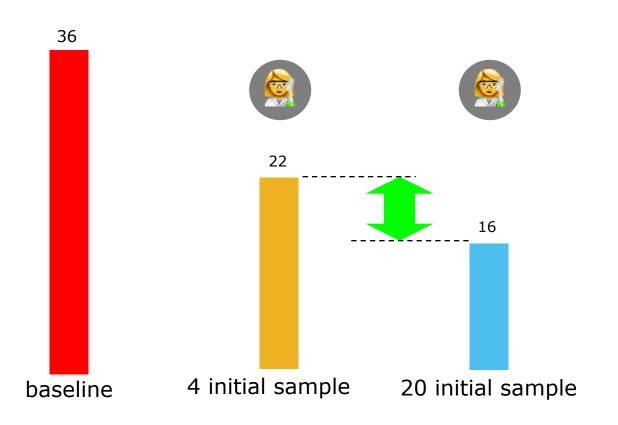
4 scenarios *

25 randomized runs =

900 Experiments/Algorithm

Q2: Training data





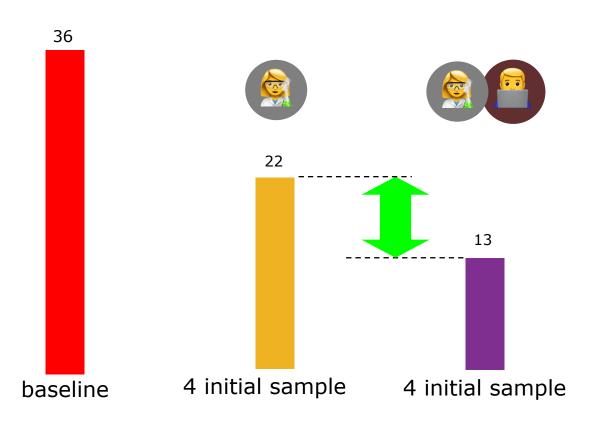
Mean required development cycle to reach top 5% performance with 90% certainty.

9 data sets *
25 randomized runs *
6 algorithms =

1350 Experiments/Scenario

Q3: Naïve vs. informed





Mean required development cycle to reach top 5% performance with 90% certainty.

9 data sets *
25 randomized runs *
6 algorithms =

1350 Experiments/Scenario

Performance in different scenarios



Mean development cycle to meet target with 90% certainty

4 initial sample 20 initial sample

naive



22

16

informed



13

10

Configuration decisions an their impact



Decision	Best decision (requ.	Alternative decision	Improvement				
	dev. cycle)	(requ. dev. cycle)	(requ. dev. cycle)				
Optimization target	Multi-objective (13.8)	Single-objective (22.2)	8.4				
Model	Gauss Process (17.7) (statistic model)	Random Forrest (18.3) (Machine Learning model)	0.6				
Pipeline	Vanilla + PCA (13.7)	Vanilla (23.5)	9.8				
Strategy	Exploration (15)	Exploitation (20.9)	5.9				
Initial sample size	20 (13.9)	4 (22)	8.1				



1. Select prediction model

explores better

Statistics-based model (**Gaussian Process** regression, GP)

- when the material data is continuous and relatively low dimensional

predicts better, but requires more data

AI model (**Random Forrest** regression, RF)

- when categorical descriptors (e.g. cement type)
- input space is not continuous, i.e. it contains completely different materials

General advice: Start with GP and try out RF if predictions are not satisfying

S BAM

2. Adjust curiosity

<0 (exploit)

- when sufficient prediction data available (sufficient data + good model)
- when training data similar to candidates & models are certain
- when deadlines are pressing

0> (explore & exploit)

- to discover "moon-shot" materials (e.g. if predictions are not satisfying)
- at the beginning of experimentation and for long-term studies



3. four criteria for the selection of candidates

Criteria 1: Utility

- the higher the better
- multi-objective optimization: adjust weights to prioritize desired property combinations

Criteria 2: Prediction & uncertainty

- predicted materials properties desired?
- if exploiting: uncertainty tolerable?



3. four criteria for the selection of candidates

Criteria 3: **A-priori information** (only if applicable)

– does the candidate meet the requirements?

Criteria 4: **Novelty** (between 0 and 1)

- indicating the difference from the training data
- a relatively low novelty (e.g. 0.05) does not introduce major changes in the composition and might therefore receive a lower priority



3. four criteria for the selection of candidates

If criteria 1...4 do not provide satisfying results:

- > try different model (GP vs. RF)
- adjust weights to change prioritization
- adjust curiosity to explore more (if only undesired materials predicted) or exploit more if good candidates already available



Design Task

www.bam.de

Geopolymer concrete parameters



AUSTRALIAN JOURNAL OF CIVIL ENGINEERING, 2018 https://doi.org/10.1080/14488353.2018.1450716





A quantitative method of approach in designing the mix proportions of fly ash and GGBS-based geopolymer concrete

G. Mallikarjuna Rao^a and T. D. Gunneswara Rao^b

^aDepartment of Civil Engineering, Vardhaman College of Engineering, Hyderabad, India; ^bDepartment of Civil Engineering, National Institut of Technology, Warangal, India

ABSTRACT

This paper presents a systematic approach for selecting mix proportions for fly ash and GGBS-based geopolymer concrete. Very little information is available on complete methodology in designing the fly ash and GGBS-based geopolymer mix. The fly ash and GGBS were activated using sodium silicate and sodium hydroxide as alkaline activator solution. The Na₂SiO₃/NaOH (alkaline activator) ratio was taken as 2.5 and the concentration of NaOH solution was maintained at 8 M. The main parameters considered in this study were binder content and alkaline solution/binder ratio for various combinations of fly ash and GGBS. The variables considered in this experimentation include: binder content (360, 420 and 450 kg/m³), proportions of fly ash and GGBS (70–30, 60–40 and 50–50), alkaline solution/binder ratios (0.45, 0.50, 0.55 and 0.60) and curing condition (outdoor curing and oven curing). Results concluded that the GGBS content, alkaline solution/binder ratio and curing condition are found to be most influential parameters on compressive strength and workability of geopolymer concrete. The paper presents detailed examples of mix designs for various strengths.

ARTICLE HISTORY

Received 1 April 2017 Accepted 5 March 2018

KEYWORDS

Geopolymer concrete; alkaline activator; compressive strength; outdoor curing; mix design

Design Parameters

Fly-ash/GGBFS ratio: 70-30, 60-40, 50-50

W/C: 0.45, 0.5, 0.55, 0.6

Binder content: 360, 370, 380, 390,400,

410, 420, 430, 440, 450 kg

Curing: ambient/oven

Target: impact of the above on compressive strength

Task: Discover a powerful concrete recipe!



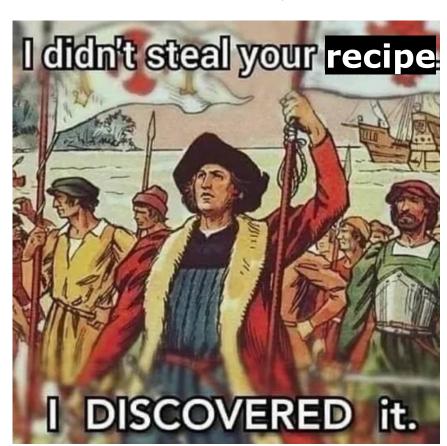
You get the instructions for a batch of 240 **geopolymer** concretes.

Step 1

 Create the formulations using SLAMD's digital twin

Step 2

 given 10 laboratory validation data use the inverse design approach to find ideal formulations







SLAMD - Sequential Learning App for Materials Discovery



Leverage the Digital Lab and AI optimization to discover exciting new concrete recipes

- > Represent resources and processes and their socio-economic impact.
- > Calculate complex compositions and enrich them with detailed material knowledge.
- > Integrate laboratory data and apply it to novel formulations.
- > Tailor concretes to the purpose to achieve the best solution.

Workflow

Step 1



1. Specify resources

From base materials to manufacturing processes - "Base" enables a detailed and consistent description of existing resources.



2. Combine resources

The combination of base materials and processes offers an almost infinite optimization potential. "Blending" makes it easier to design complex configurations.



3. Digital recipes

With "Formulations" you can effortlessly convert your resources into the entire spectrum of possible concrete formulations. This automatically generates a detailed set of data for AI optimization.

Step 2



4. Materials Discovery

Integrate data from the "Digital Lab" or upload your own material data. Enrich the data with lab results and adopt the knowledge to new recipes via artificial intelligence. Leverage socio-economic metrics to identify recipes tailored to your requirements.

Link: https://slamddemo.herokua pp.com



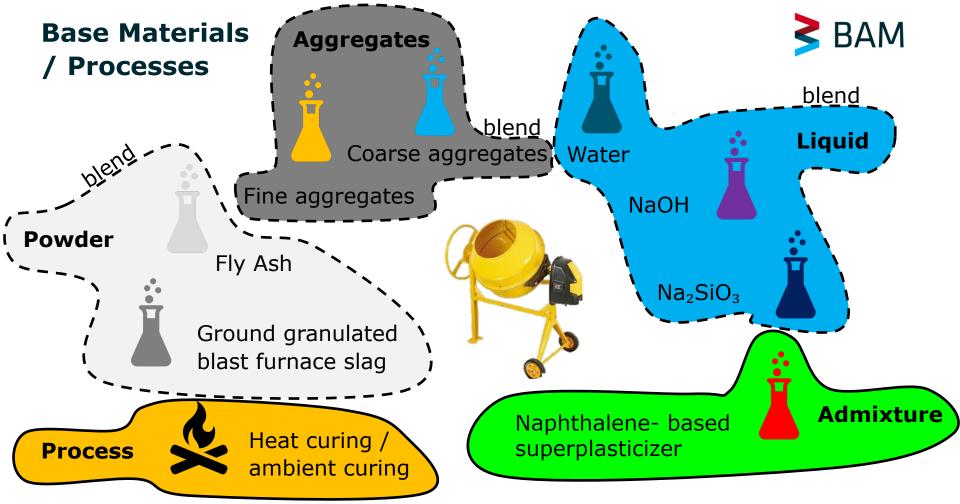
Links











Base Materials - Powder



		Composit	ion								Cost
<u>:</u>		Fine modulus (m²/kg)	Fe ₂ O ₃ (m%)	SiO ₂ (m%)	Al ₂ O ₃ (m%)	CaO (m%)	MgO (m%)	Na₂O (m%)	SO ₃ (m%)	LOI (m%)	CO ₂ footprint (kg/ton)
T	FA	380	4.25	60.11	26.53	4.0	1.25	0.22	0.35	3.25	4
Ë	GGBFS *	426	0.8	34.06	20.0	32.6	7.89	0.0	0.9	3.72	52.0

The example session file already contains some materials. Materials with star "*"must be added in SLAMD.

Base Materials - Liquid



		Compositio	Composition							
		Na ₂ SiO ₃ (m%)	NaOH (m%)	H ₂ O (m%)	CO ₂ footprint (kg/ton)					
*	Water	0	0	100	0					
Å	NaOH*	0	100	0	1915.0					
	Na ₂ SiO ₃	100	0	0	360.0					

The example session file already contains some materials. Materials with star "*"must be added in SLAMD.

Base Materials - Aggregates





	Composit	tion					Cost
	Fine Aggregates (m%)	Coarse Aggregates (m%)	Specific gravity (kg/m³)	Bulk density (kg/m³)	Fineness modulus (m³/kg)	Water absorption (m%)	CO ₂ footprint (kg/ton)
Fine	100	0	2.65	1.45	2.57	2.0	4.8
Coarse	0	100	2.8	1.5	7.3	0.5	4.8

Base Materials - Admixture



Cost

CO₂ footprint (kg/ton)



Sulphonated Naphthalene formaldehydebased superplasticizer *

1880

The example session file already contains some materials. Materials with star "*"must be added in SLAMD.

Process - Curing



	Process Inform	mation	Cost
	Duration (days)	Temperature (°C)	CO ₂ footprint (kg)
Heat*	1	60	22.45
Ambient	1	25	0

The example session file already contains some materials. Materials with star "*"must be added in SLAMD.

Blends



```
Powders (3): FA / GGBFS
```

50 / 50 60 / 40 70 / 30

Liquid (1): $H_2O / NaOH / Na_2SiO_3$

72 / 8 / 20

Aggregates (1): Fine / Coarse

40 / 60

Formulation



(# of variations) Design Parameters

- (3) Fly-ash/GGBFS ratio: 70-30, 60-40, 50-50
- (4) Alkaline solution/binder ratio: 0.45, 0.5, 0.55, 0.6
- (1) Admixture: Super plasticiser 4 wt. % of Binder
- (2) Curing: ambient/oven
- (10) Binder content: 360, 370, 380, 390,400, 410, 420, 430, 440, 450 kg
- (1) Fine/coarse aggregate: 45 wt.% / 55 wt.%

3 x 4 x 1 x 2 x 10 -> **240** Formulations

Constraint (Sum of materials used for formulation) = 2400 kg $\sim 1m^3$

Design Tasks



1. Single-objective

Target = high strength $f_{c(28d) > 61.94}MPa \quad (5\% -> 12/240)$

2. Multi-objective

Target = high strength & climate friendly

$$f_{c(28d)} > 50 MPa$$

 $CO_2 < 85 kg/m^3 \quad (5\% -> 12/240)$

	Blend	Curing	w/c				Powe	der Co	ntent -	- (kg)				
Hands on!	8	3	>	360	370	380	390	400	410	420	430	440	450	
		4	0,45	55,4	55,6	55,9	56,1	56,4	56,6	56,9	54,1	51,9	48,9	
		ien	0,5	59,8	59,9	60,0	60,1	60,2	60,3	60,4	59,8	59,2	58,5	
		Ambient	0,55	51,6	50,7	49,9	49,0	48,1	47,3	46,5	47,2	47,7	48,5	
	50/50	•	0,6	46,7	46,8	47,0	47,1	47,3	47,4	47,6	47,6	47,6	47,6	
Ground Truth Data	20'		0,45	62,0	61,2	60,5	59,7	59,0	58,2	57,5	56,7	55,8	55,0	_
Ground math bata		Heat	0,5	65,3	65,1	64,9	64,8	64,6	64,4	64,3	63,5	62,7	61,9	Target
		Ξ̈́	0,55	53,4	52,9	52,5	52,0	51,5	51,1	50,6	51,9	53,2	54,4	
			0,6	52,4	51,9	51,4	50,9	50,4	49,9	49,4	50,9	52,4	53,8	64 MPa
		t	0,45	43,3	43,0	43,6	44,2	44,9	45,5	46,1	46,0	46,0	45,9	
		ien	0,5	47,9	48,5	47,9	48,8	49,6	50,4	51,2	51,0	50,9	50,7	
		Ambient	0,55	46,9	46,5	46,1	45,8	45,4	45,0	44,7	43,7	42,9	41,9	
	60/40		0,6	43,4	46,2	46,3	46,4	46,4	46,5	46,6	46,3	46,0	45,7	
	09			0,45	55,6	56,1	56,7	57,2	57,8	58,3	58,8	58,0	57,3	56,6
		Heat	0,5	57,4	58,2	59,0	59,9	60,7					61,3	47 MPa
		Ξ̈́	0,55	53,4	54,3	55,3	56,2	57,1	58,1	58,9	57,4	55,8	54,4	
			0,6	51,2	52,1	53,1	54,0	54,9	55,8	56,6	54,3	52,0	49,8	
		٠	0,45	33,8	34,3	34,8	38,3	35,8	36,2	36,7	35,7	34,9	33,8	
		Ambient	0,5	36,2	36,5	36,9	37,2	37,5	37,9	38,2	38,4	38,7	39,0	
		\mb	0,55	31,1	30,3	29,4	28,6	27,8	27,0	26,2	26,0	25,9	25,7	
	70/30	4	0,6	25,7	25,6	25,5	25,4	25,3	25,2	25,1	23,6	22,4	20,8	
	70/		0,45	41,5	41,3	41,2	41,0	40,8	40,6	40,4	39,8	39,2	38,6	20 MPa
		at	0,5	42,6	42,7	42,9	43,0	43,2	43,4	43,5	43,8	44,2	44,5	20 1917 a
		Heat	0,55	33,9	33,5	33,0	32,6	32,1	31,7	31,2	32,6	34,0	35,2	
25.01.24			0,6	25,6	28,3	28,0	27,7	27,4	27,2	26,9	27,8	28,7	29,5	39

Questions?





"SLAMD" - Sequential Learning App for Materials Discovery Repository: https://github.com/BAMresearch/WEBSLAMD

Resources:

https://github.com/BAMresearch/WEBSLAMD#documentation

- Video Tutorial
- Benchmarking Paper
- Data Sets
- GPT-4 powered SLAMD Assistant