Automate engineering processes at scale

Efficient optimization of material properties based on experimental data

User conference 2023

Johannes Gerritzen (TUD), Nina Moëllo (Datadvance)
Dr. Andreas Hornig (TUD), Laurent Chec (Datadvance)
Prof. Dr. Maik Gude (TUD)

Content

- Presentation of the company
- Introduction to the optimization problem
- Proposed approach
- Methodology
- Conclusion and perspectives

Technische Universität Dresden

- **17 faculties** in **5 areas** (Engineering sciences, Humanities and social sciences, Mathematics and natural sciences, Civil and environmental engineering, Medicine)
- More than **30 000 students**, thereof about 50 % in engineering
- More than **8 200 employees** → one of the biggest employers in SME dominated saxony
- One of Germans top ranked **Excellence Universities** since 2012
- Dynamic, cosmopolitan, family-friendly, strong in research since **foundation in 1828**

VISION

The ILK is the leading international institute for research, development in the field of function-integrative lightweight engineering based on multi-material design

ILK-TEAM

245

employees in a broad inter-disciplinary team

70

years of lightweight research in Dresden

YOUNG TALENTS

80

graduates per year

CONTINUOUS RESEARCH AND DEVELOPMENT CHAINS

Material, Modelling, Simulation, Design, Processing, Quality, LCA

INDUSTRIAL COOPERATION

with European large-scale industry and regional SMEs

~1.000

Alumni since 1997

>1

Start-Up a year (at present 18)

35%	Basic research
35%	Application-oriented research
30%	Industrial Development

Promoting initiatives for children, women and young talent

ACL e.V.

juniorING e.V.

INTERNATIONAL NETWORK

among others with partners in UK, Poland, Korea, China, Singapore, Romania, Australia, USA

Starting point

- Limited availability of materials meeting requirements of future jet engines
- Development of novel material
  - Extremely expensive experiments
  - Influence of controllable variables unclear
  - Properties only partially accessible to simulations

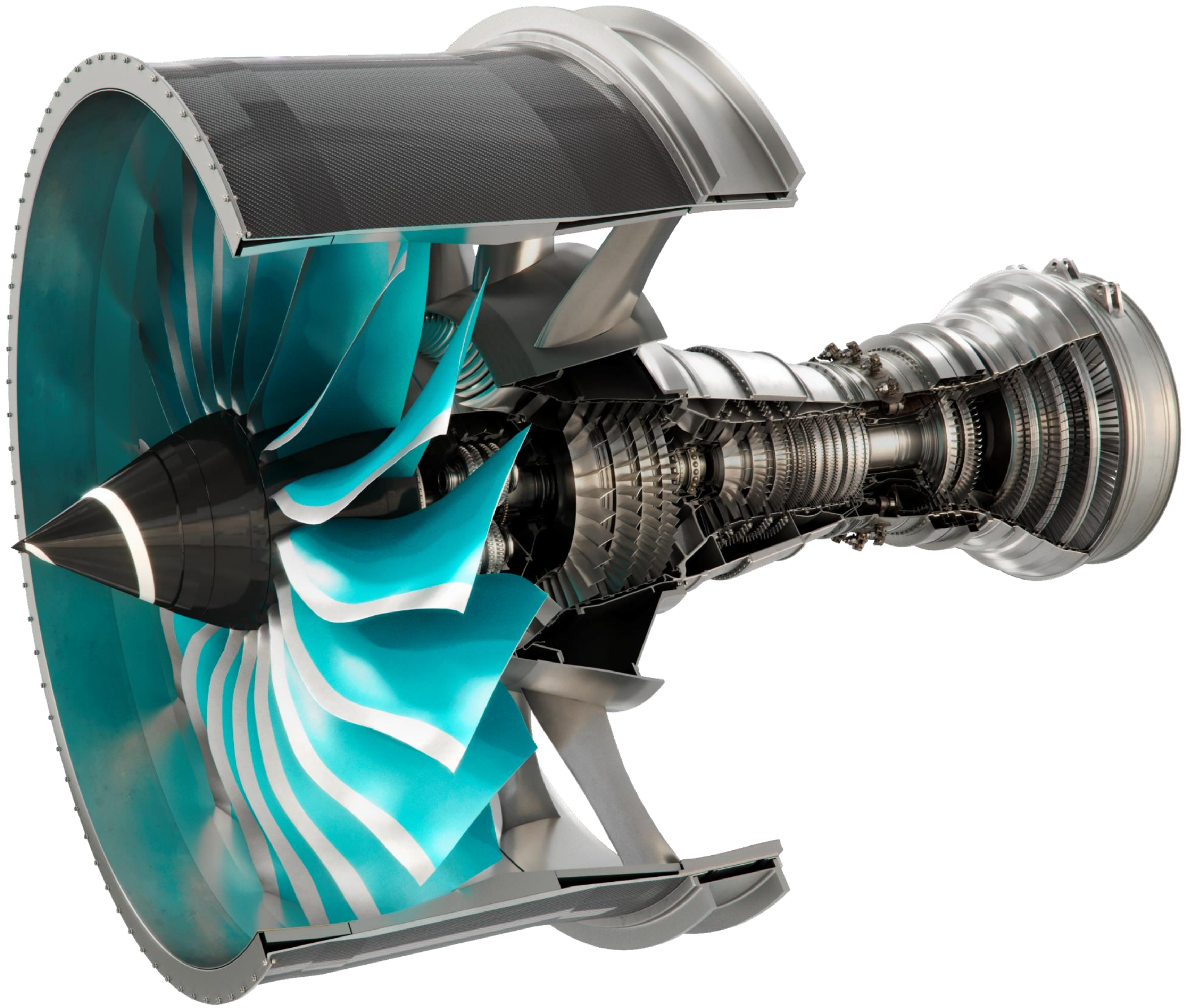


Image source: https://www.flickr.com/photos/rolls-royceplc/48486636831/in/album-72157644584413758/#

Introduction to the optimization problem

- **Experiment based** analysis to study **properties of materials**.
- The goal is to guide the experiments in order to **optimize the resulting material's performance**.

10 inputs

Chemical composition
Pressures
Temperatures
Process timing

→

2 objectives and 3 constraints

Density ↘
Young's modulus ↗
Elongation
Yield strength
Ultimate tensile strength (UTS)

Introduction to the optimization problem

How to extract the optimal point?

- Limitations due to the **high dimensionality** of input domain and **multi-objective** optimization, while having a **reduced budget**.
- Limited behavioral information: the database contains **only 35 points,** sometimes with **incomplete information**. The ideal number of points is 10 x (number of variables) = 10 x 10 = 100 points to get a relevant model.
- **No** access to a **simulation tool** to run as blackbox.

Using only experimental approach to get to the optimal design means a lot of tests, which takes a very long time.

Proposed approach

- **Step 1**: populate the database to obtain as much information as possible. Using approximation models, the missing values are predicted.

	A	B	C
1	Point 1	Point 2	Point 3
2	1.5	2	?
3	3.8	4.9	0.9
4	12	?	0.46

BEST ?

- **Step 2**: use the database to give the optimal point. This part relies on surrogate-based optimization (SBO).

- Two workflows were created to automate the process.

Methodology – step 1

- First database: **10 points** were given with **all the output values** information.
- Second database: **25 points** were given with **only the Young's modulus** output values.

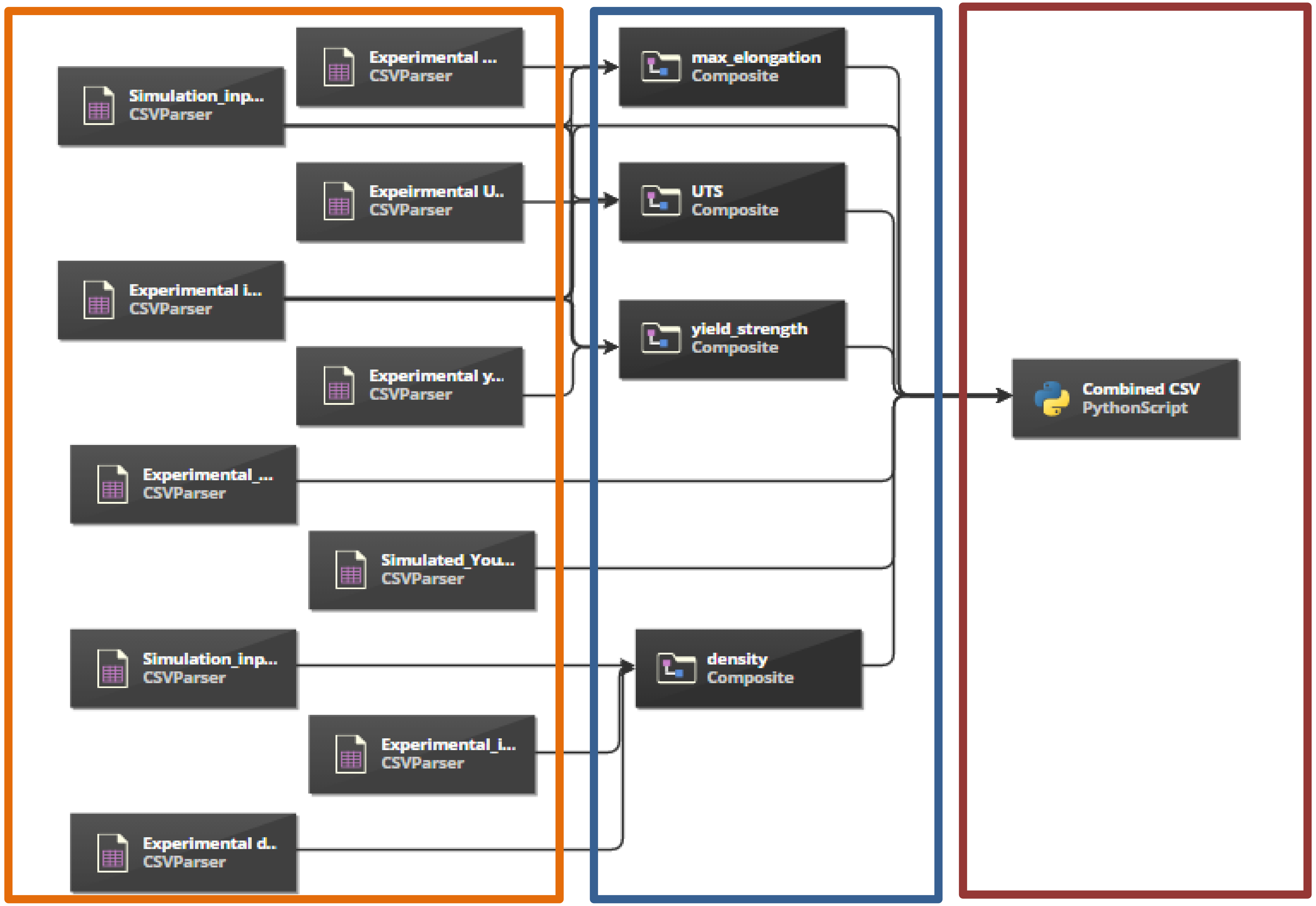
- Based on the first database, **approximation models were created** for the density, elongation, UTS and yield strength.
- These models were then used to **predict missing outputs** of the second database.
- In the end, both databases are merged.

Methodology – step 1

1. Load CSV databases.

2. Create models and evaluate missing values.

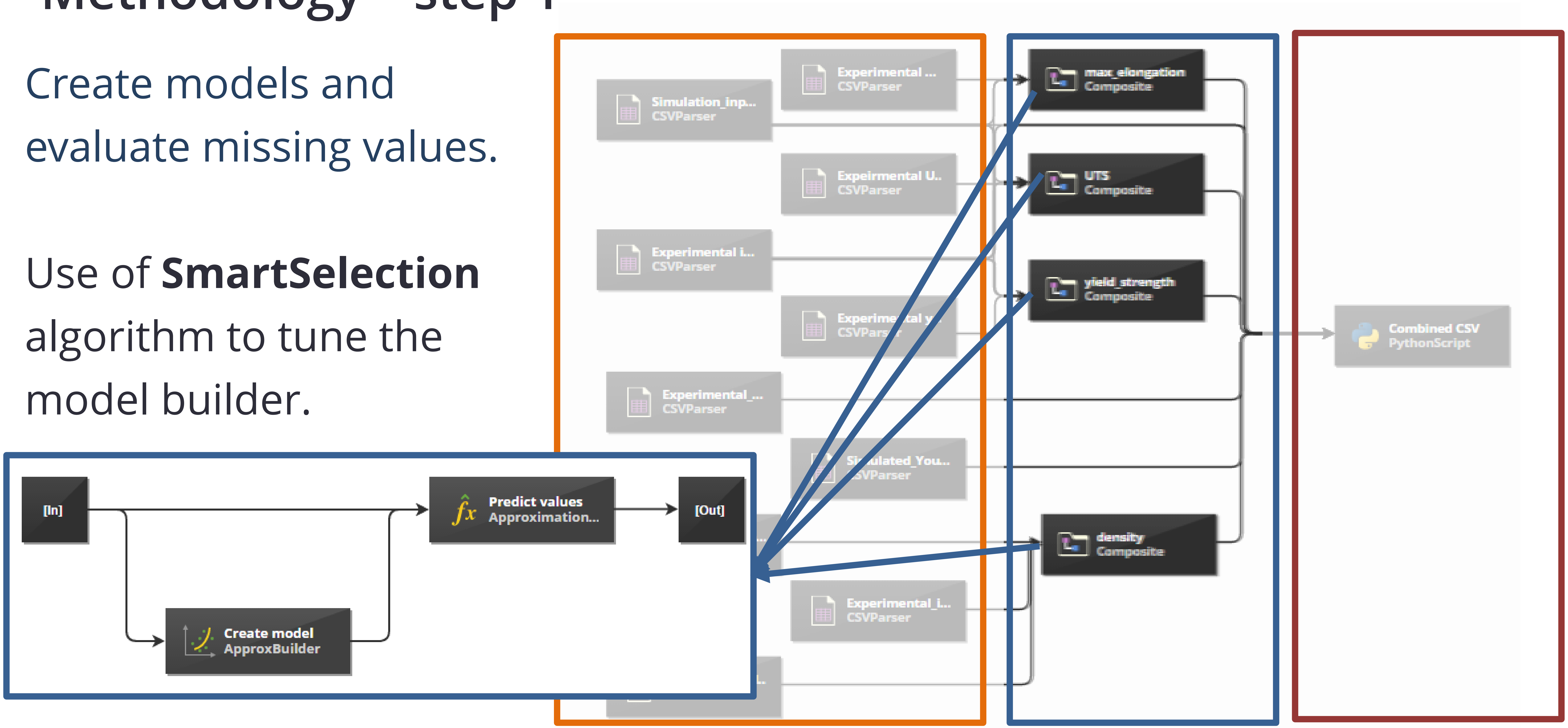
3. Merge all into one CSV file.



Methodology – step 1

Create models and evaluate missing values.

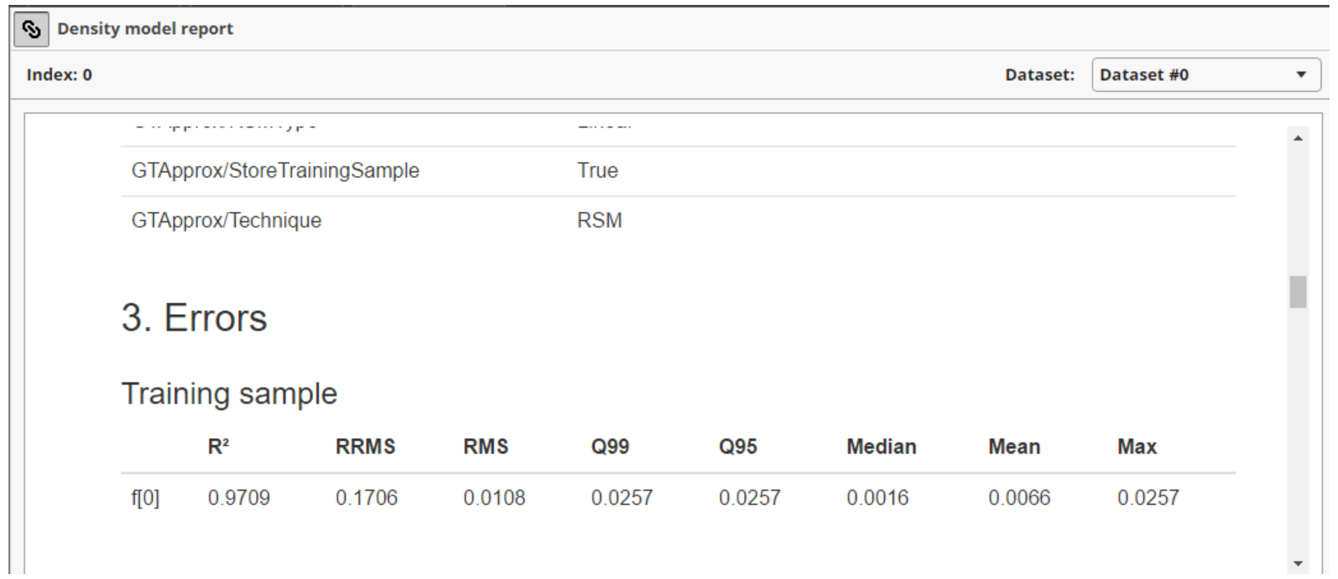
Use of **SmartSelection** algorithm to tune the model builder.



Methodology – step 1

- **Automated process** to create models and evaluate missing values. Can be **easily launched** to populate again the database **after an update** (new experiment).

- **SmartSelection** algorithm allows fine tuning of the model builders to create **performant approximations without the involvement of an expert**. All the models are **independent** and can be checked using **generated reports** if any doubt.



Density model report

Index: 0 Dataset: Dataset #0

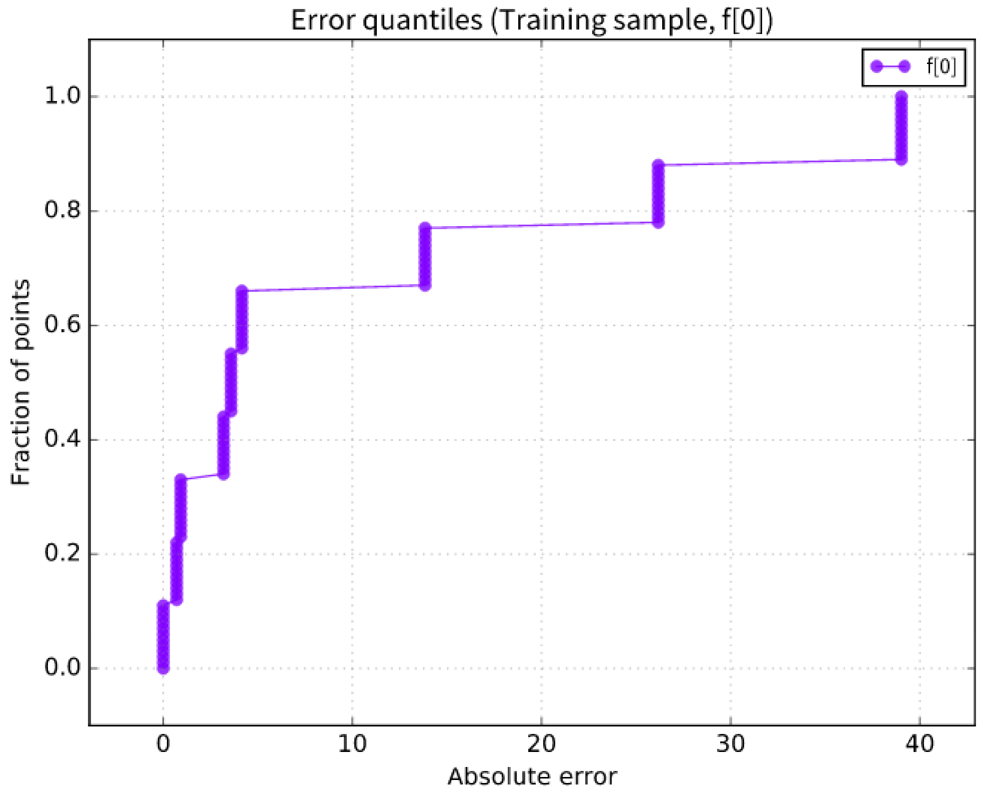
GTApprox/StoreTrainingSample	True
GTApprox/Technique	RSM

3. Errors

Training sample

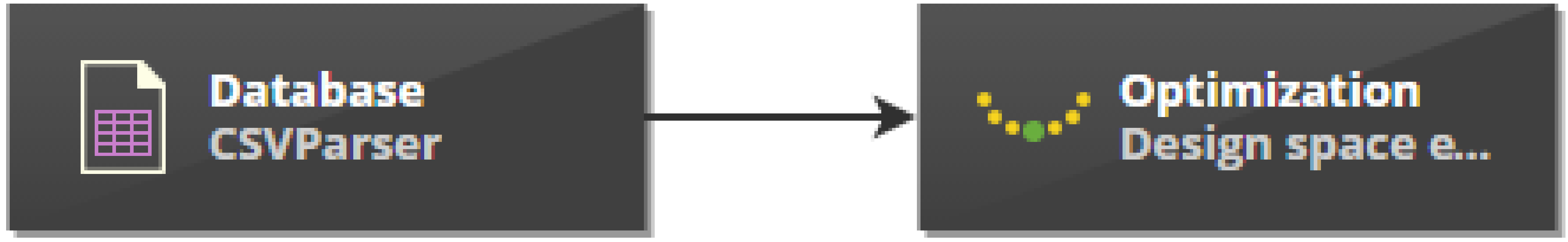
	R^2	RRMS	RMS	Q99	Q95	Median	Mean	Max
f[0]	0.9709	0.1706	0.0108	0.0257	0.0257	0.0016	0.0066	0.0257

3. Error quantiles - training sample



Methodology – step 2

- Once the **database is complete**, it is used in a **second workflow** for optimization.
- A **Design Space Exploration** block manages the **optimization**, taking as information input the database.



Methodology – step 2

- **SmartSelection: SBO**

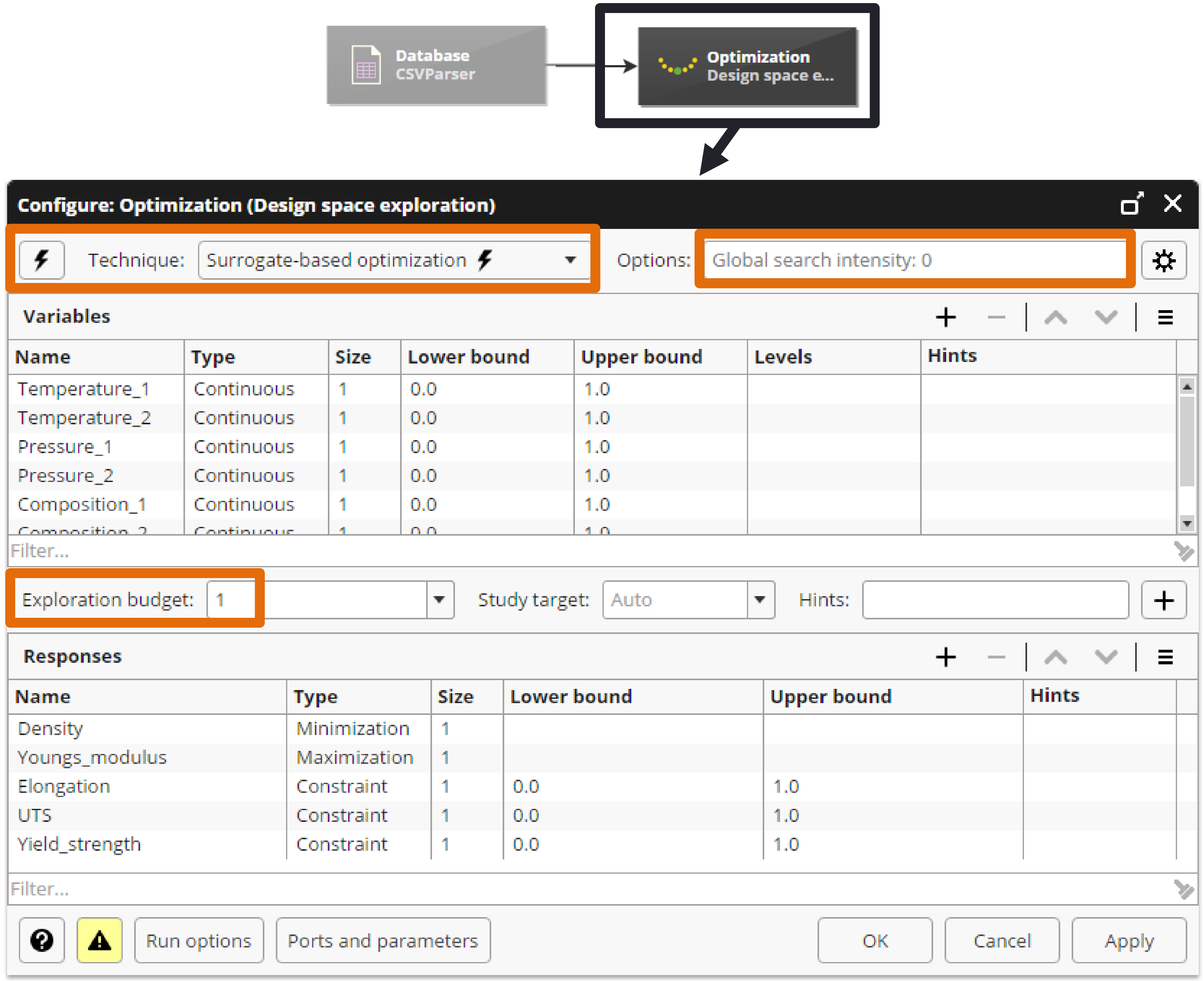
  Based on the given points, a surrogate model is created, and the best point of this model is selected as optimal design point.

- **Exploration budget: 1**

  As there is no iterative process, the software is asked to provide directly the best point out of the surrogate model.

- **Global search intensity: 0**

  No budget allowed to explore the design space, only optimization.



NB : the names and values of Variables and Responses are fake for confidentiality purpose.

Conclusion and perspectives

- An **automated approach** to build models and populate the databases when needed. The pSeven process takes around **2 minutes**.
- Before optimization, a first check of the models and predictions is possible through **automatically generated reports**.
- The optimizer is parametrized to give the **optimal point** at the **first computation**, relying on a dedicated **surrogate model**.
- Even with **low data** (i.e., 1/3 of what is required for such operation), the models can give a **tendency**.
- **Reduced complexity and time** compared to only experimental testing.

NB : both initial databases are from different sources, thus different fidelity but because of the very low amount of data, they have been gathered in the same file for the optimization. If more data have been available, the Data Fusion approach would be selected.

THANK YOU

Visit us:

- pseven.io

Contact us:

- info@pseven.io
- 42 Avenue du Général de Croutte, 31100, Toulouse, FRANCE
- Tel: +33 (0) 6 03-84-62-92

Follow us:

- youtube.com/datadvance
- linkedin.com/company/datadvance
- facebook.com/datadvance

Johannes Gerritzen (TUD), Nina Moëllo (Datadvance)
Dr. Andreas Hornig (TUD), Laurent Chec (Datadvance)
Prof. Dr. Maik Gude (TUD)

funded by
Bundesministerium
für Wirtschaft
und Klimaschutz

