Supporting Information

Automating and improving GHG emissions calculation in pharma / fine chemicals synthesis R&D: GreenSpeed as a digital tool to navigate complex value chains

Alexander Dauth*, Benjamin Kühne*, Helmut Hänsel, Sara Wirsing, Kerstin Hell, Heinrich Becker, Isabelle Georg, Andreas Bathe, Bertram Cezanne

Data Availability

GreenSpeed was coded using The Julia Programming Language (https://julialang.org/) and is hosted on internal servers at Merck KGaA.

The ELNs used at Merck KGaA and connected to GreenSpeed are Revvity Signals (https://revvitysignals.com/) and ELAB by enso Software (https://www.enso-software.com/)

Merck KGaA's Synthia® can be accessed through https://www.synthiaonline.com/

Merck KGaA's DOZN™ can be accessed through https://www.sigmaaldrich.com/DE/en/services/software-and-digital-platforms/dozn-tool

All other tools, standards, legal texts and white papers can be accessed as referenced in the main article

GreenSpeed Core features:

The landing page is kept intentionally lean, only requiring the input of experiment IDs and the choice of the source ELN.

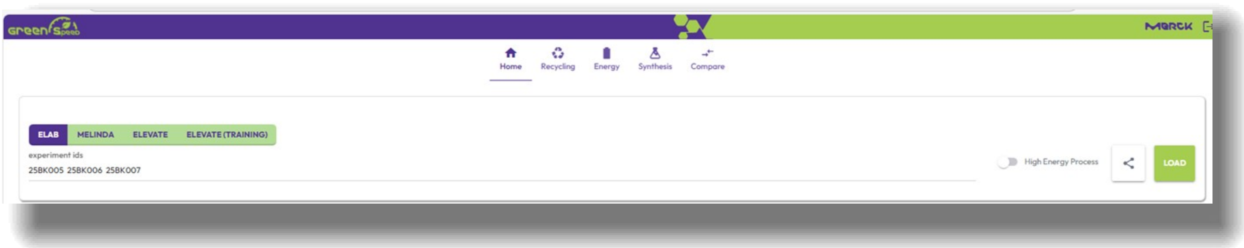


Figure 1. GreenSpeed landing page – only experiment IDs from ELN need to be entered (e.g. via copy & paste).

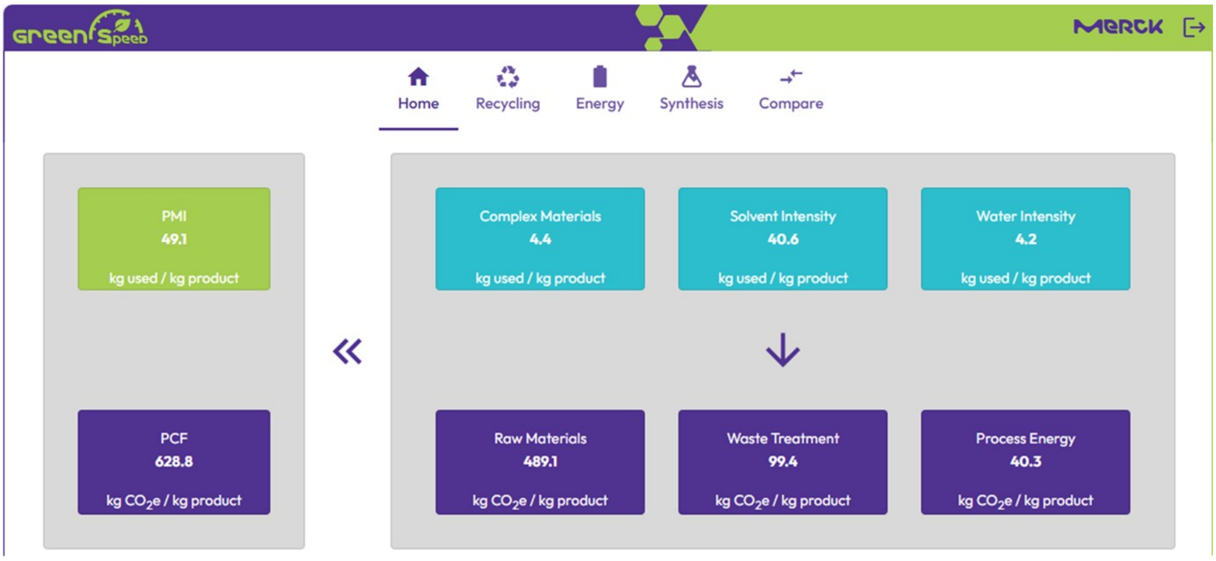


Figure 2. PMI and PCF overview screen.

The most important KPIs are printed in the boxes on top of the page as shown in Figure 1.

$$\text{Process Mass Intensity (PMI)} = \frac{\sum \text{mass of materials}}{\text{mass of isolated product}}$$

$$\text{Solvent Intensity (SI)} = \frac{\sum \text{mass of used solvents}}{\text{mass of isolated product}}$$

$$\text{Water Intensity (WI)} = \frac{\sum \text{mass of water in the process}}{\text{mass of isolated product}}$$

Complex Materials = PMI – SI - WI

The colors of the box for PMI give a qualitative assessment of the greenness of the process with regard to resource consumption (green: good, avg. step PMI < 20, yellow: ok, avg. step PMI < 40, red: bad, avg. step PMI > 40). This relies on experience values from standard organic synthesis processes for complex small molecules in organic solvent and might not be applicable for more exotic processes.

The Product Carbon Footprint (PCF, estimated total CO_2 release, Cradle-to-Grave) is approximated as the sum of three contributions: Raw Materials CF + Waste Treatment + Process Energy.

Raw Materials CF is calculated from carbon footprint factors for each material. If no factor can be found in either the expert list or proxy list, it defaults to role-based factors. Alternatively, a custom carbon factor can be entered manually (see also What-If Scenarios – Modify Carbon Footprint)

1. **Expert List**: This data source includes a curated selection of carbon footprint data derived from supplier information and secondary databases combined with expert assessments. The entries in this list are limited, focusing on the most relevant and reliable data sources to ensure a high accuracy in the PCF calculations. Most common solvents and frequently used raw materials are covered here.

2. **Proxy List**: The proxy list was developed using a combination of carbon footprint mapping methodologies that consider both spend-based and weight-based approaches. It has a wider coverage, but lower data quality compared to the Expert List.

3. **Role-Based Classification**: This classification organizes raw materials based on their roles in the synthesis process. It includes the following categories: reactant / chemical building block (100 kg CO_2e/kg), reagent (5 kg CO_2e/kg), catalyst (1000 kg CO_2e/kg), solvent (4.9 kg CO_2e/kg), and water (0.001 kg CO_2e/kg). This is the fall-back option if no data is found for chemicals via the other sources and defaults to pre-set values for the applied roles. The values chosen for the various roles were based on rounded averages for the different categories. The catalyst footprints relate to precious metals and include recycling rates (often >90 %).

The **Waste Treatment** contribution is based on Complex Materials, SI and WI and assigns a factor for the treatment of each group. Waste CF = Solvent Intensity * 2.3 + Water Intensity * 0.63 + Complex Materials * 0.78 [kg per kg product]

Process Energy can be approximated using the energy tab and otherwise defaults to a correlation to the PMI. The option "High Energy Process" can be chosen for a different correlation factor considering processes with an unusually high energy demand (e.g. calcinations or sublimations):
Regular processes: 0.82 (kg CO2/kg Product) * PMI
High energy processes: 2.45 * PMI
These correlations are based on calculations and averages from almost 40 internal organic production processes from all of Merck KGaA's business sectors on manufacturing scale.

The table shown in Figure 2 lists the same KPIs as defined above.

The top rows ("cumulated/final") show cumulated KPIs, these are the KPIs per 1 kg of the product of this step including the steps before.

The bottom rows ("step") present the KPIs for the single steps. This can be used this to identify the step which should be most urgently optimized.

Experiment ID	Mode	PMI (without Cleaning)	PMI	Carbon Footprint (cradle-to-gate) [kg/kg prod.] [1]	Solvent-Intensity	Water-Intensity	Complex-Materials	Raw Materials CF	Process Energy [3]	Waste Treatment [2]
[illegible]	FINAL STEP (cumulated)	49.1	49.1	628.8	40.6	4.2	4.4	489.1	40.3	99.4
[illegible]	cumulated	33.6	33.6	451.6	25.9	4.2	3.6	359	27.6	65
[illegible]	cumulated	13.3	13.3	282.5	7.2	3.6	2.6	250.9	10.9	20.8
[illegible]	step	16.7	16.7	281.6	14.9	0	1.9	232.2	13.7	35.7
[illegible]	step	19.2	19.2	240	17.5	0	1.7	182.6	15.8	41.6
[illegible]	step	13.3	13.3	282.5	7.2	3.6	2.6	250.9	10.9	20.8

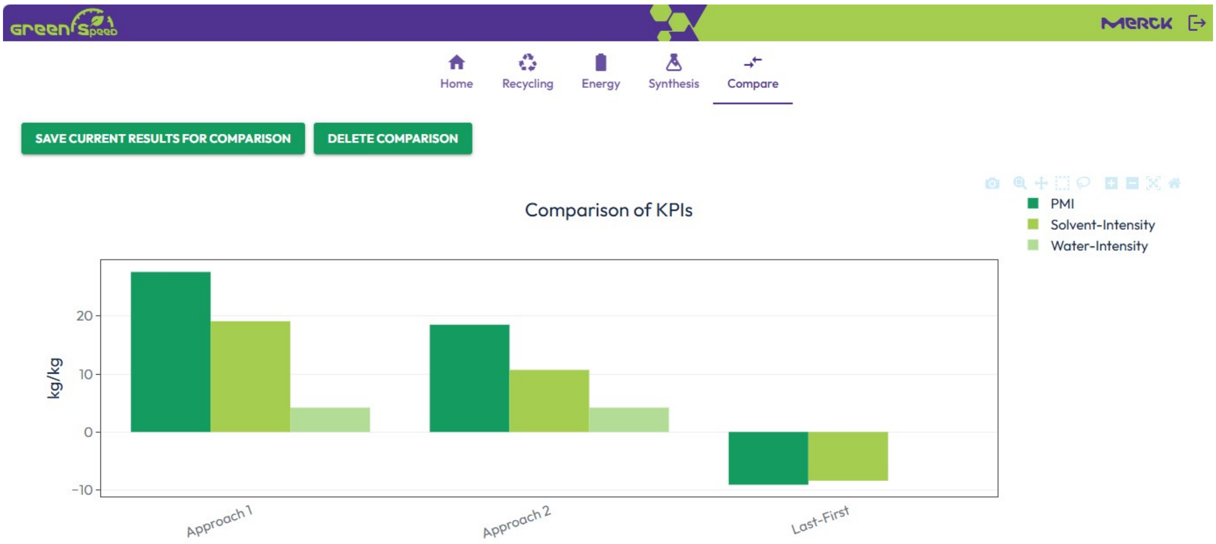
Records per page: All 1-6 of 6

DOWNLOAD KPIS

Figure 3. PMI and PCF evaluation table

Extended features:

Direct Comparison: Two or more synthesis routes can be directly compared and the difference in PMI, PCF and other metrics can be displayed graphically and in table-form (Figure 3). This facilitates identification and quantification of the route with the lowest environmental impact among several candidates. Often, this feature is also used to show the improvement of the same synthesis after process development and optimization.



IDs	PMI	Carbon Footprint (cradle-to-gate) [kg/kg prod.] [1]	Solvent-Intensity	Water-Intensity	Complex-Materials	Raw Materials CF	Process Energy [3]	Waste Treatment [2]
elab:25BK003,elab:25BK002,elab:25BK001	27.6	507.8	19.1	4.2	4.4	435.2	22.6	49.9
elab:25BK003,elab:25BK002,elab:25BK001	18.5	422.7	10.7	4.2	3.7	377.6	15.2	30
Last-First	-9.1	-85.1	-8.4	0	-0.7	-57.6	-7.4	-19.9

Figure 4. Direct comparison of two synthesis routes.

What-if Scenarios: Users can modify material quantities (1) and raw material footprints (2) to evaluate the implications of different choices on sustainability metrics. The source of the carbon footprint is shown (3) to allow an assessment of the overall data quality. The effect of e.g. omitting a work-up step or substituting a raw material with a low PCF alternative can be evaluated (Figure 4).

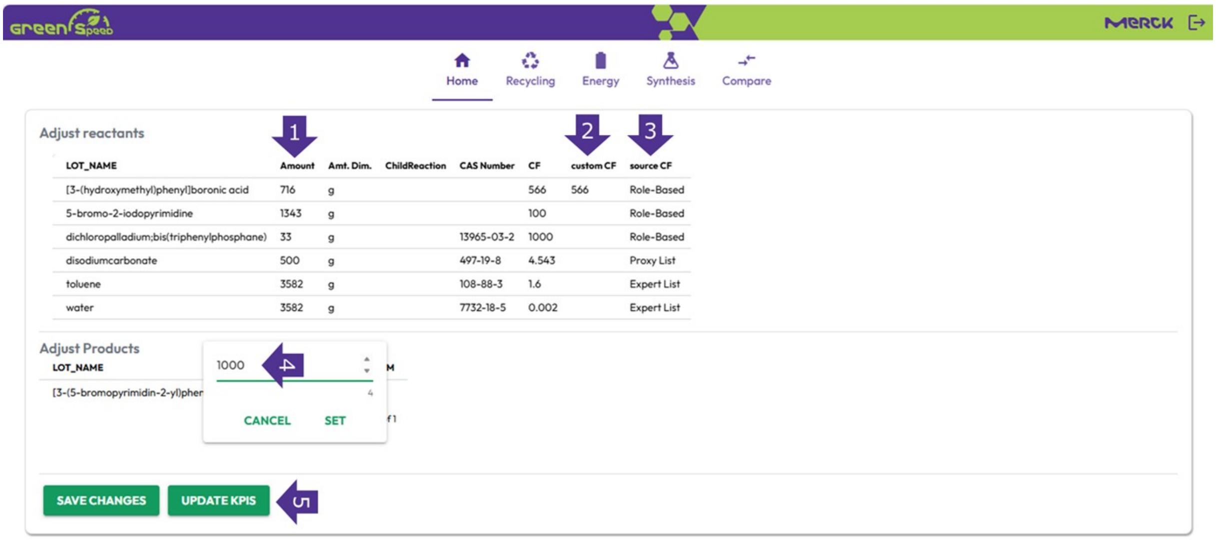


Figure 5. Adjusting quantities and carbon footprints manually allows for scenario modeling.

Similarity Search: This feature enables researchers to identify similar transformations among all of Merck KGaA's ELN entries, providing yield vs PMI plots that visualize potential improvements for any organic synthesis step that has been documented in the ELN. Through this, expected targets can be set for new reactions and previous work can inform new approaches.

The degree of similarity can be adjusted through a slider to retrieve more focused or wider reaching results (Figure 5).

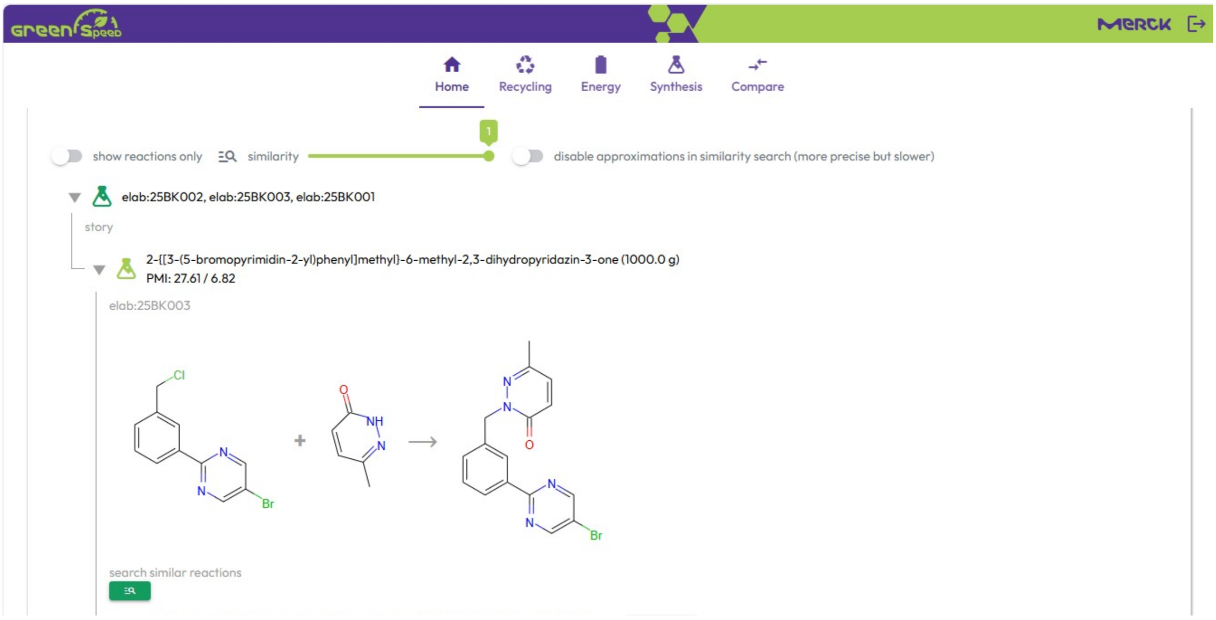


Figure 6. Similarity search can be performed on individual steps, selecting the degree of similarity.

The experiments found are displayed in tabular form and in a Yield vs. PMI graph (Figure 6) which plots PMI and yield of all simlar entries retrieved from the ELN. This can be used to compare an ELN entry of interest to previous work. The details of the entries with lower PMI/higher yield can be used to improve a given process. The color code differentiates between entries made for lab work vs. pilot plant. The latter entries are often of notably higher data quality.

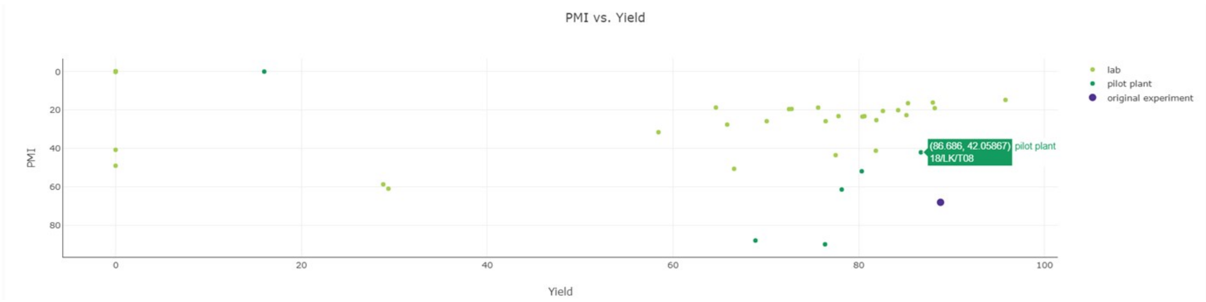


Figure 7. Yield vs PMI plot of similar experiments.

Recycling Assessment: The application evaluates the recycling potential of materials used in synthesis, promoting practices aligned with the principles of circular economy. Individual waste streams can be marked for recycling and the effect on the waste treatment footprint is displayed (Figure 7). This enables scientists to not only develop processes but also consider by-products and devise recycling routines.

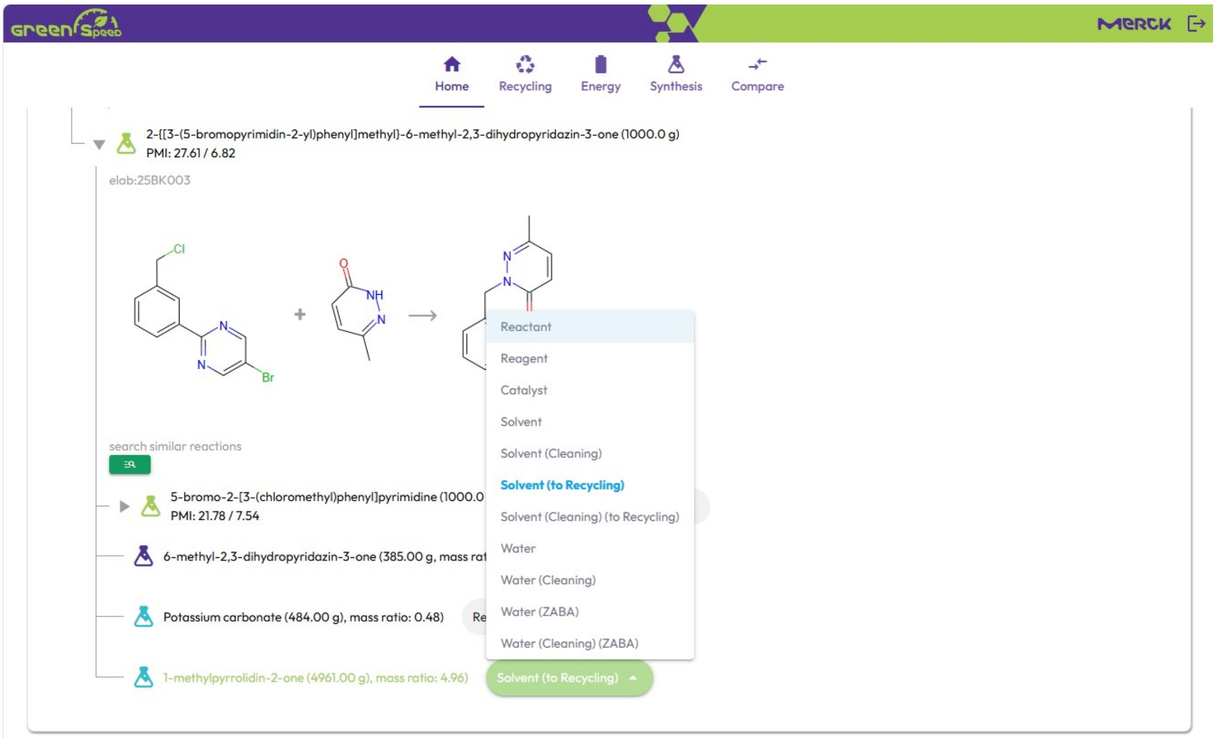
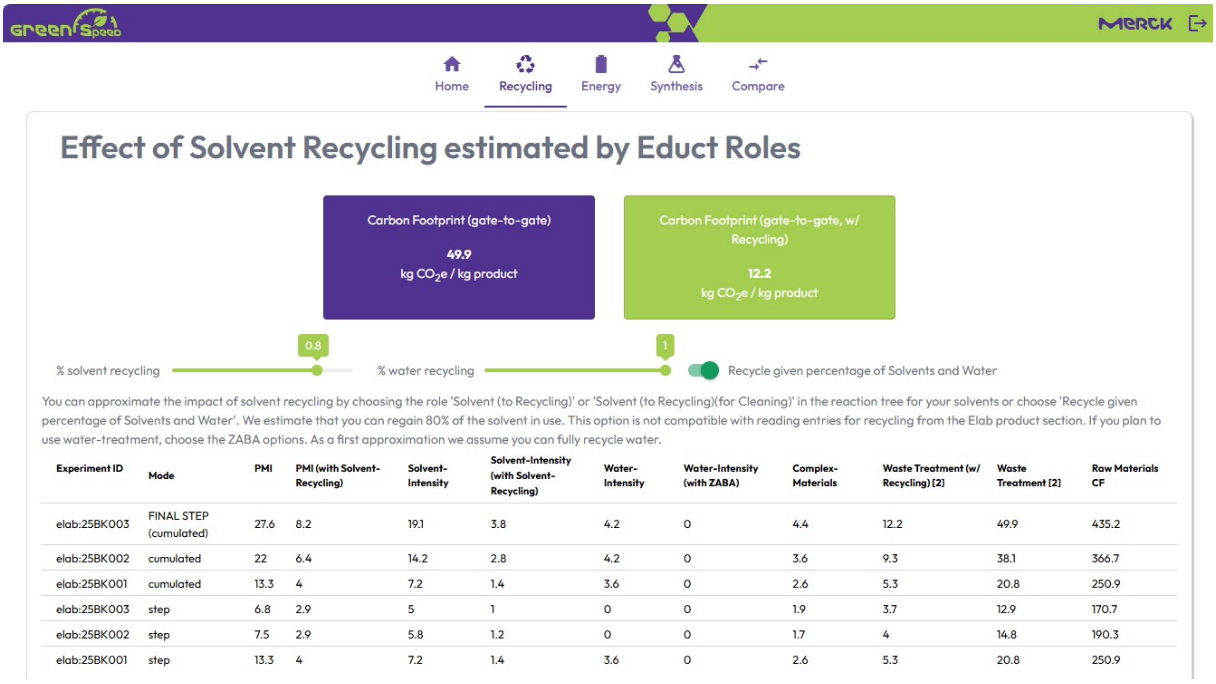


Figure 8. Selection of individual solvent streams for recycling.

The recycling percentage defaults to the estimate that 80% of solvent can be gained back (Figure 8). This includes actual loss from incomplete recycling but also approximately accounts for the energy which needs to be employed to recycle the solvents. Water is approximated to be fully transferred back after treatment. It is also possible to individually define the percentages of the (selected) water and solvent streams to be recycled.



Experiment ID	Mode	PMI	PMI (with Solvent-Recycling)	Solvent-Intensity	Solvent-Intensity (with Solvent-Recycling)	Water-Intensity	Water-Intensity (with ZABA)	Complex-Materials	Waste Treatment (w/ Recycling) [2]	Waste Treatment [2]	Raw Materials CF
elab:25BK003	FINAL STEP (cumulated)	27.6	8.2	19.1	3.8	4.2	0	4.4	12.2	49.9	435.2
elab:25BK002	cumulated	22	6.4	14.2	2.8	4.2	0	3.6	9.3	38.1	366.7
elab:25BK001	cumulated	13.3	4	7.2	1.4	3.6	0	2.6	5.3	20.8	250.9
elab:25BK003	step	6.8	2.9	5	1	0	0	1.9	3.7	12.9	170.7
elab:25BK002	step	7.5	2.9	5.8	1.2	0	0	1.7	4	14.8	190.3
elab:25BK001	step	13.3	4	7.2	1.4	3.6	0	2.6	5.3	20.8	250.9

Figure 9. Potential GHG emissions reduction through solvent recycling.

GreenSpeed PMI and PCF calculation results for starting materials 1, 2, and 3:

Detailed breakdown of PMI and PCF to the main contributors is shown. The life cycle inventory for each synthesis is based on the modeling principles and chemical equations shown in Scheme 1 in the publication.

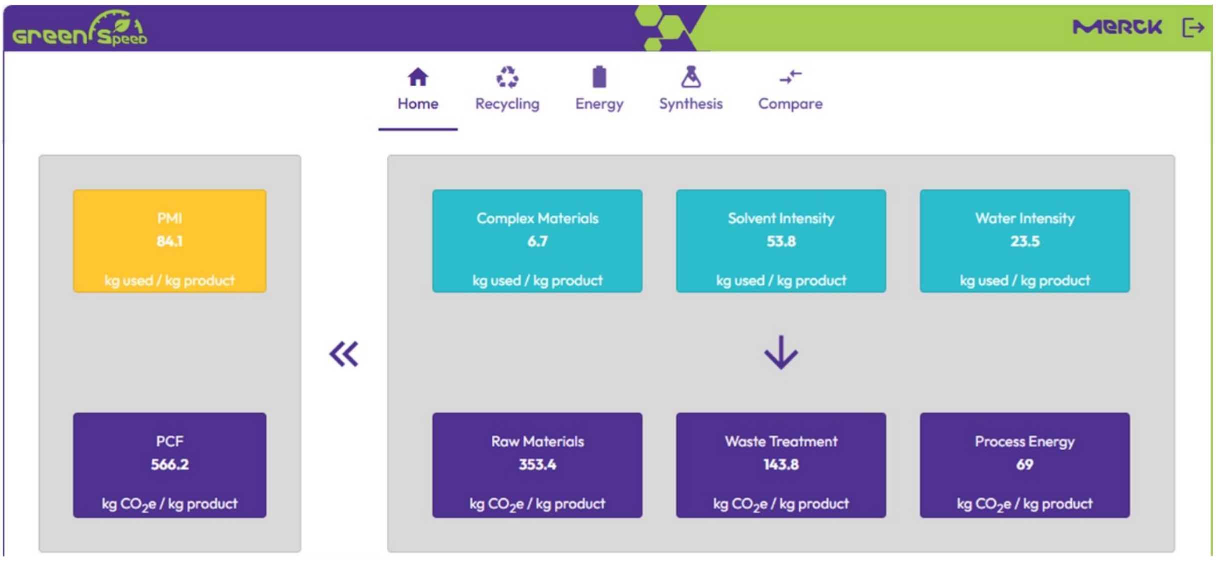


Figure 10. PMI and PCF calculation for starting material 1.

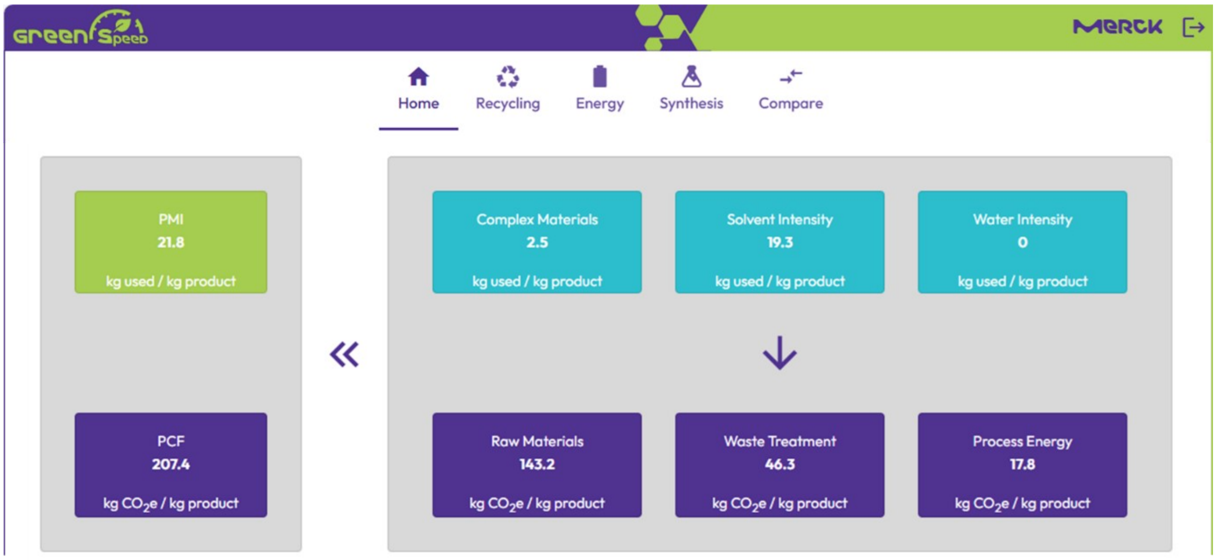


Figure 11. PMI and PCF calculation for starting material 2.

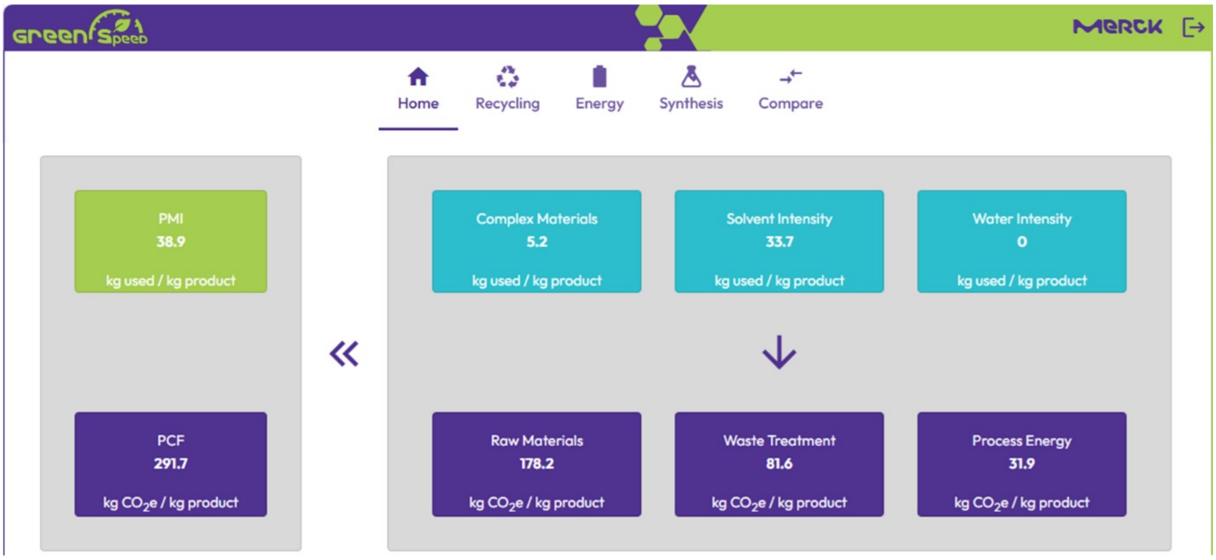
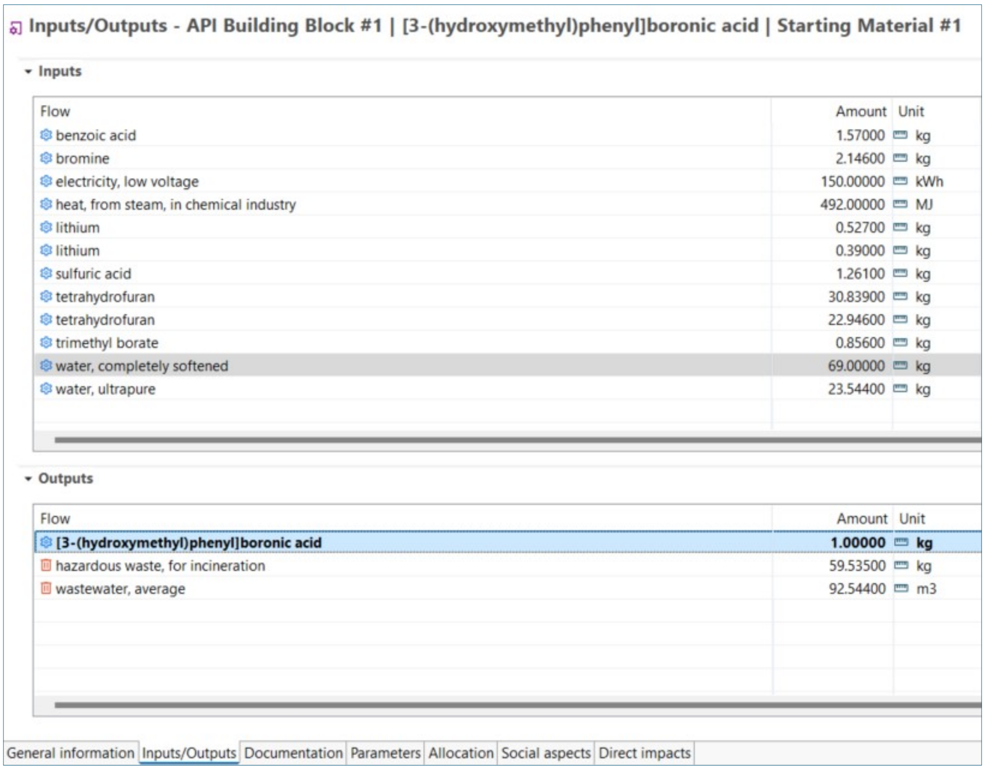


Figure 12. PMI and PCF calculation for starting material 3.

OpenLCA calculation results for starting materials 1, 2, and 3:



Inputs/Outputs - API Building Block #1 | [3-(hydroxymethyl)phenyl]boronic acid | Starting Material #1

Inputs

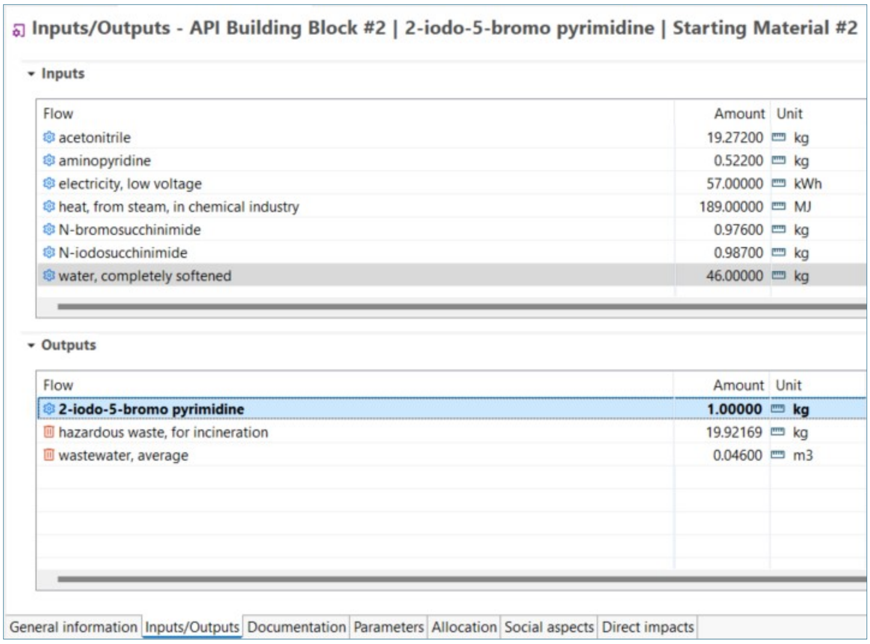
Flow	Amount	Unit
benzoic acid	1.57000	kg
bromine	2.14600	kg
electricity, low voltage	150.00000	kWh
heat, from steam, in chemical industry	492.00000	MJ
lithium	0.52700	kg
lithium	0.39000	kg
sulfuric acid	1.26100	kg
tetrahydrofuran	30.83900	kg
tetrahydrofuran	22.94600	kg
trimethyl borate	0.85600	kg
water, completely softened	69.00000	kg
water, ultrapure	23.54400	kg

Outputs

Flow	Amount	Unit
[3-(hydroxymethyl)phenyl]boronic acid	**1.00000**	**kg**
hazardous waste, for incineration	59.53500	kg
wastewater, average	92.54400	m3

General information | Inputs/Outputs | Documentation | Parameters | Allocation | Social aspects | Direct impacts

Figure 13. Life Cycle Inventory (LCI) in OpenLCA to evaluate environmental footprints for starting material 1, used to benchmark GreenSpeed calculations as shown in Table 1 in the paper.



Inputs/Outputs - API Building Block #2 | 2-iodo-5-bromo pyrimidine | Starting Material #2

Inputs

Flow	Amount	Unit
acetonitrile	19.27200	kg
aminopyridine	0.52200	kg
electricity, low voltage	57.00000	kWh
heat, from steam, in chemical industry	189.00000	MJ
N-bromosuccinimide	0.97600	kg
N-iodosuccinimide	0.98700	kg
water, completely softened	46.00000	kg

Outputs

Flow	Amount	Unit
2-iodo-5-bromo pyrimidine	**1.00000**	**kg**
hazardous waste, for incineration	19.92169	kg
wastewater, average	0.04600	m3

General information | Inputs/Outputs | Documentation | Parameters | Allocation | Social aspects | Direct impacts

Figure 14. Life Cycle Inventory (LCI) in OpenLCA to evaluate environmental footprints for starting material 2, used to benchmark GreenSpeed calculations as shown in Table 1 in the paper.

Inputs/Outputs - API Building Block #3 | Chloro maleic hydrazide | Starting Material #3

Inputs

Flow	Amount	Unit
acetic anhydride	1.22200	kg
electricity, low voltage	68.00000	kWh
heat, from steam, in chemical industry	222.00000	MJ
hydrazine	0.38400	kg
maleic anhydride	1.17400	kg
N-methyl-2-pyrrolidone	16.10100	kg
phosphorus oxychloride	1.46800	kg
sodium chloroacetate	0.98200	kg
triethyl amine	17.60700	kg
water, completely softened	46.00000	kg

Outputs

Flow	Amount	Unit
Chloro maleic anhydride	**1.00000**	**kg**
hazardous waste, for incineration	36.53200	kg
wastewater, average	0.04600	m3

General information | Inputs/Outputs | Documentation | Parameters | Allocation | Social aspects | Direct impacts

Figure 15. Life Cycle Inventory (LCI) in OpenLCA to evaluate environmental footprints for starting material 3, used to benchmark GreenSpeed calculations as shown in Table 1 in the paper.

ACS Streamlined PMI-LCA tool calculation results for starting materials 1, 2, and 3:

Step Name	Input or Output	LCA Data Source Class	LCA Data Source Subclass	Display Name	Physical Mass (kg)
Step 1	Input	iGAL_Aligned_Organic	Default iGal Aligned Organic Reagent	benzoic acid	1,57
Step 1	Input	Common_Inorganic	Default Inorganic Material	potassium bromate	2,15
Step 1	Input	Common_Inorganic	Default Inorganic Material	sulfuric acid	1,26
Step 1	Input	Common_Solvent	water, deionized	water	23,55
Step 1	Output	None	None	3-Bromobenzoic acid	2,06
Step 2	Input	Process_Steps	Step 1	3-Bromobenzoic acid	2,06
Step 2	Input	Common_Inorganic	Default Inorganic Material	lithium aluminium hydride	0,39
Step 2	Input	Common_Solvent	tetrahydrofuran (THF)	THF	30,84
Step 2	Output	None	None	(3-bromophenyl)methanol	1,53
Step 3	Input	Process_Steps	Step 2	(3-bromophenyl)methanol	1,53
Step 3	Input	iGAL_Aligned_Organic	Default iGal Aligned Organic Reagent	butyl lithium	0,53
Step 3	Input	Common_Inorganic	Default Inorganic Material	methyl borate	0,86
Step 3	Input	Common_Solvent	tetrahydrofuran (THF)	THF	22,95
Step 3	Output	None	None	[3-(hydroxymethyl)phenyl]boronic acid	1,00
PACKAGE	Input	Process_Steps	Step 3	[3-(hydroxymethyl)phenyl]boronic acid	1,00
PACKAGE	Output	None	None	Packaged [3-(hydroxymethyl)phenyl]boronic acid	1,00

Process Metrics per kg API	Total	Reagent	Metal	Solvent	Water
PMI	84,1	6,8	0,0	53,8	23,6
Mass Net (kg)	209,3	10,6	0,0	198,7	0,0
Energy (MJ)	7008,6	206,5	0,0	6801,7	0,4
GWP (kg CO2 equiv.)	319,6	11,9	0,0	307,7	0,0
Acidification (kg SO2 equiv.)	1,1	0,1	0,0	1,0	0,0
Eutrophication (kg posphate equiv.)	0,4	0,0	0,0	0,4	0,0
Water (kg)	1146,1	34,0	0,0	1082,3	29,8
COG (USD)	0,0	0,0	0,0	0,0	0,0

Figure 16. Life Cycle Inventory (LCI) and results in Streamlined PMI-LCA tool for starting material 1 synthesis, used to benchmark GreenSpeed calculations as shown in Table 1 in the paper.

Step Name	Input or Output	LCA Data Source Class	LCA Data Source Subclass	Display Name	Physical Mass (kg)
Step 1	Input	Non_iGAL_Aligned_Orga	Default Non-iGAL Aligned Organic Reagent	2-amino pyrimidine	0,52
Step 1	Input	iGAL_Aligned_Organic	Default iGal Aligned Organic Reagent	N-bromosucchinimide	0,98
Step 1	Input	Common_Solvent	acetonitrile	acetonitrile	7,82
Step 1	Output	None	None	2-amino-5-bromo pyrimidine	0,76
Step 2	Input	Process_Steps	Step 1	2-amino-5-bromo pyrimidine	0,76
Step 2	Input	iGAL_Aligned_Organic	Default iGal Aligned Organic Reagent	N-iodosucchinimide	0,99
Step 2	Input	Common_Solvent	acetonitrile	acetonitrile	11,45
Step 2	Output	None	None	2-iodo-5-bromo pyrimidine	1,00
PACKAGE	Input	Process_Steps	Step 2	2-iodo-5-bromo pyrimidine	1,00
PACKAGE	Output	None	None	Packaged 2-iodo-5-bromo pyrimidine	1,00

Process Metrics per kg API	Total	Reagent	Metal	Solvent	Water
PMI	21,8	2,5	0,0	19,3	0,0
Mass Net (kg)	68,5	29,3	0,0	39,2	0,0
Energy (MJ)	2319,8	649,1	0,0	1670,7	0,0
GWP (kg CO2 equiv.)	114,7	55,9	0,0	58,7	0,0
Acidification (kg SO2 equiv.)	0,7	0,4	0,0	0,3	0,0
Eutrophication (kg posphate equiv.)	0,2	0,1	0,0	0,2	0,0
Water (kg)	267,4	37,3	0,0	230,2	0,0
COG (USD)	0,0	0,0	0,0	0,0	0,0

Figure 17. Life Cycle Inventory (LCI) and results in Streamlined PMI-LCA tool for starting material 2 synthesis, used to benchmark GreenSpeed calculations as shown in Table 1 in the paper.

Step Name	Input or Output	LCA Data Source Class	LCA Data Source Subclass	Display Name	Physical Mass (kg)
Step 1	Input	iGAL_Aligned_Organic	Default iGal Aligned Organic Reagent	Maleic anhydride	1,17
Step 1	Input	iGAL_Aligned_Organic	Default iGal Aligned Organic Reagent	Hydrazine	0,38
Step 1	Input	Common_Solvent	triethanolamine (TEA)	triethylamine	17,61
Step 1	Input	iGAL_Aligned_Organic	Default iGal Aligned Organic Reagent	acetic anhydride	1,22
Step 1	Input	iGAL_Aligned_Organic	Default iGal Aligned Organic Reagent	sodium acetate	0,98
Step 1	Output	None	None	Maleic hydrazide	1,07
Step 2	Input	Process_Steps	Step 1	Maleic hydrazide	1,07
Step 2	Input	Common_Inorganic	Default Inorganic Material	phosphoryl chloride	1,47
Step 2	Input	Common_Solvent	1-methyl-2-pyrrolidinone/N-methyl pyrrol	1-methylpyrrolidin-2-one	16,10
Step 2	Output	None	None	Chloro maleic anhydride	1,00
PACKAGE	Input	Process_Steps	Step 2	Chloro maleic anhydride	1,00
PACKAGE	Output	None	None	Packaged Chloro maleic anhydride	1,00

Process Metrics per kg API	Total	Reagent	Metal	Solvent	Water
PMI	38,9	5,2	0,0	33,7	0,0
Mass Net (kg)	97,3	9,8	0,0	87,5	0,0
Energy (MJ)	3454,0	297,3	0,0	3156,7	0,0
GWP (kg CO2 equiv.)	141,3	11,9	0,0	129,4	0,0
Acidification (kg SO2 equiv.)	0,5	0,1	0,0	0,5	0,0
Eutrophication (kg posphate equiv.)	0,9	0,0	0,0	0,9	0,0
Water (kg)	487,0	56,3	0,0	430,7	0,0
COG (USD)	0,0	0,0	0,0	0,0	0,0

Figure 18. Life Cycle Inventory (LCI) and results in Streamlined PMI-LCA tool for starting material 3 synthesis, used to benchmark GreenSpeed calculations as shown in Table 1 in the paper.