nature chemical engineering

Article

https://doi.org/10.1038/s44286-024-00033-5

Autonomous reaction Pareto-front mapping with a self-driving catalysis laboratory

In the format provided by the authors and unedited

Table of Contents

1. Chemicals

All chemicals were used as received. Toluene (Anhydrous 99.8%) and tridecane analytical standard were purchased from MilliporeSigma. 1-Octene, 99+%, and triphenylphosphine, 99%, were purchased from Thermo Scientific Chemicals. Acetylacetonato dicarbonyl rhodium(I) ($Rh(CO)_2(acac)$), 97%, was purchased from Alfa Aesar. XantPhos, 98%, was purchased from Strem Chemicals. 2-octene (mixture of cis + trans isomer) (+98%) was purchased from TCI. Carbon monoxide, hydrogen, and nitrogen were purchased from Airgas at 99.9% purity. 1-octene and 2-octene were purified over silica gel columns before use. Fluorophosphite ligands L1-L3 were provided by Eastman Chemical. Synthesis details and spectra of ligands L1-L3 are reported in prior work.[1]

2. Experimental Setup

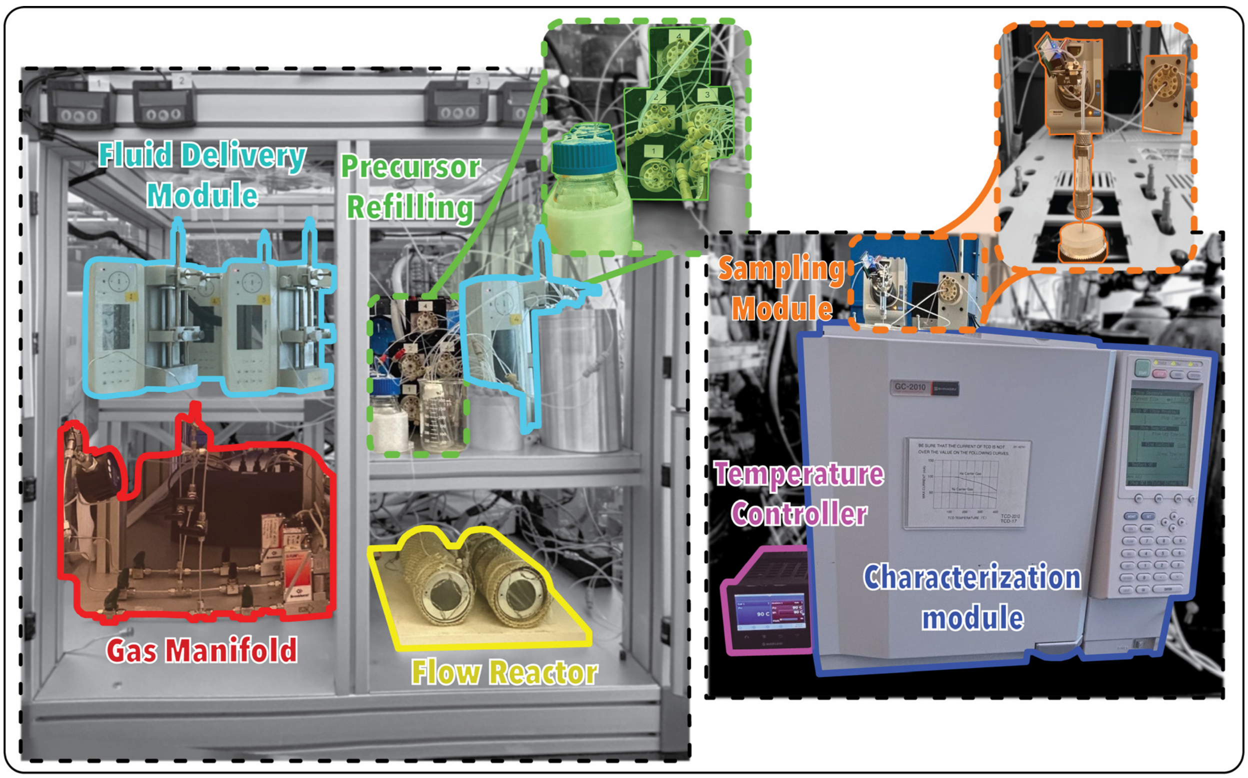

Figure 1. *Fast-Cat's experimental setup: Blue) Liquid Delivery Module. Red) Reactive Gas Manifold. Green) Precursor Refilling Module. Yellow) Heated Coil Flow Reactor. Pink) Reactor Temperature Controller. Orange) In-Line Sampling Module. Purple) GC-FID Characterization Module.*

The automated flow chemistry platform of Fast-Cat (**Fig. 1**, **Fig. 1** in the main text) is comprised of six modules, including precursor loading/formulation, precursor refilling, gas delivery, flow reactor, reaction sampling, and in-line characterization using gas chromatography (GC-FID). The flow reactor is supplied with four separate stock solutions to adjust concentrations of the individual reagent species independently and match the next predicted hydroformylation condition selected by the ML module of Fast-Cat. The liquid feeds are the olefin stream (olefin + solvent + internal standard), ligand stream (ligand + solvent + internal standard), catalyst stream (rhodium salt + solvent + internal standard), and solvent stream (solvent + internal standard). The stock solutions are prepared such that an equal volumetric flowrate results in 0.5M 1-octene, 0.25 mM $Rh(CO)_2(acac)$, 2.5 mM Ligand, and 20 mM tridecane internal standard in toluene as the solvent (**Table S1**).

Table 1. *Composition of liquid feed solutions.*

Feed	**Composition**
Olefin	2.0M 1-octene, 20mM tridecane in toluene
Ligand	10 mM ligand, 20 mM tridecane in toluene
Rhodium	1 mM $Rh(CO)_2(acac)$, 20 mM tridecane in toluene
Solvent	20 mM tridecane in toluene

After the solutions are prepared, they are loaded in the automated precursor refill reservoirs and high-pressure syringe pumps (Chemyx Fusion 6000) using 20 mL stainless steel syringes (Chemyx) and automated 10-way flow switching valves (VICI). Next, the flow reactor line is brought up to the target pressure with a 1:1 mixture of carbon monoxide (CO) and hydrogen (H_2). The process control module of Fast-Cat then takes over the operational workflow and the closed-loop catalysis cycles begin.

The hardware of Fast-Cat automatically adjusts the liquid feed composition by varying the volumetric flowrate of the individual liquid input streams, subject to volumetric ratios that maintain stable flow through the mixing junctions. The flow reactor consists of a length of 1/16" outer diameter (OD) × 0.03" inner diameter (ID) stainless steel tubing wrapped around a grooved aluminum heating block (**Fig. 1**).

The syngas composition is automatically tuned by varying individual CO and H_2 mass flow controllers (Bronkhorst MFCs) to set the flowrate and control for any supply pressure fluctuations. The reaction temperature is controlled using PID temperature controllers (Watlow F4T) with integrated thermocouples and cartridge heaters (Watlow) inside the aluminum reactor block that holds the stainless-steel reactor coil. The coil-wrapped reactor block is then insulated to minimize any temperature gradients in the flow reactor coil as well as the power consumed by the cartridge heaters.

The total pressure in the flow reactor section is controlled by a digital pressure regulator (Equilibar EHP DPR) supplying pressurized nitrogen to the control surface of a diaphragm-based backpressure regulator (Equilibar Zero Flow BPR). The BPR then holds the upstream flow reactor line above the pressure supplied to the control surface. The gas flowrates at a given pressure setpoint are calculated and used to set the total liquid volumetric flowrate to maintain a 3:1 gas:liquid volumetric ratio within the flow reactor module of Fast-Cat, ensuring the reaction in the liquid slugs is not limited by the supply of syngas in the gas phase. The total volumetric flowrate of the liquid and the gas phases defines the residence time (*i.e.*, hydroformylation reaction time) of the reactive liquid slugs within the heated flow reactor section prior to in-line sampling by the automated GC unit (Shimadzu GC-2010) equipped with a flame ionization detector. The reaction is allowed to equilibrate for four residence times to achieve steady-state operation at the set reaction condition before sampling. The automated sampling system switches a sample-loop valve to begin filling by the two-phase reactive line and an attached phase-sensor detects when the sample loop is completely filled with liquid to switch the captured sample onto the GC analysis line, push it into the GC using carrier gas (helium), and trigger the GC oven method and data acquisition. The sample loop and GC injection needle are washed between experimental conditions by automatically sending droplets of acetone through the injector line and conducting a wash and GC column bake method.

Prior to performing a new reaction condition, Fast-Cat evaluates the current available precursor volume of the syringe pumps and refills them from the precursor refilling module if the new reaction conditions would completely consume the precursors available in the stainless-steel syringes before completing the next hydroformylation reaction. The process automation module of Fast-Cat will safely shutdown operation upon reaching the allowed experimental budget and will pause gas and liquid stream flowrates if waiting on a new condition, while still under the experimental budget.

3. Automated Precursor Refilling Module

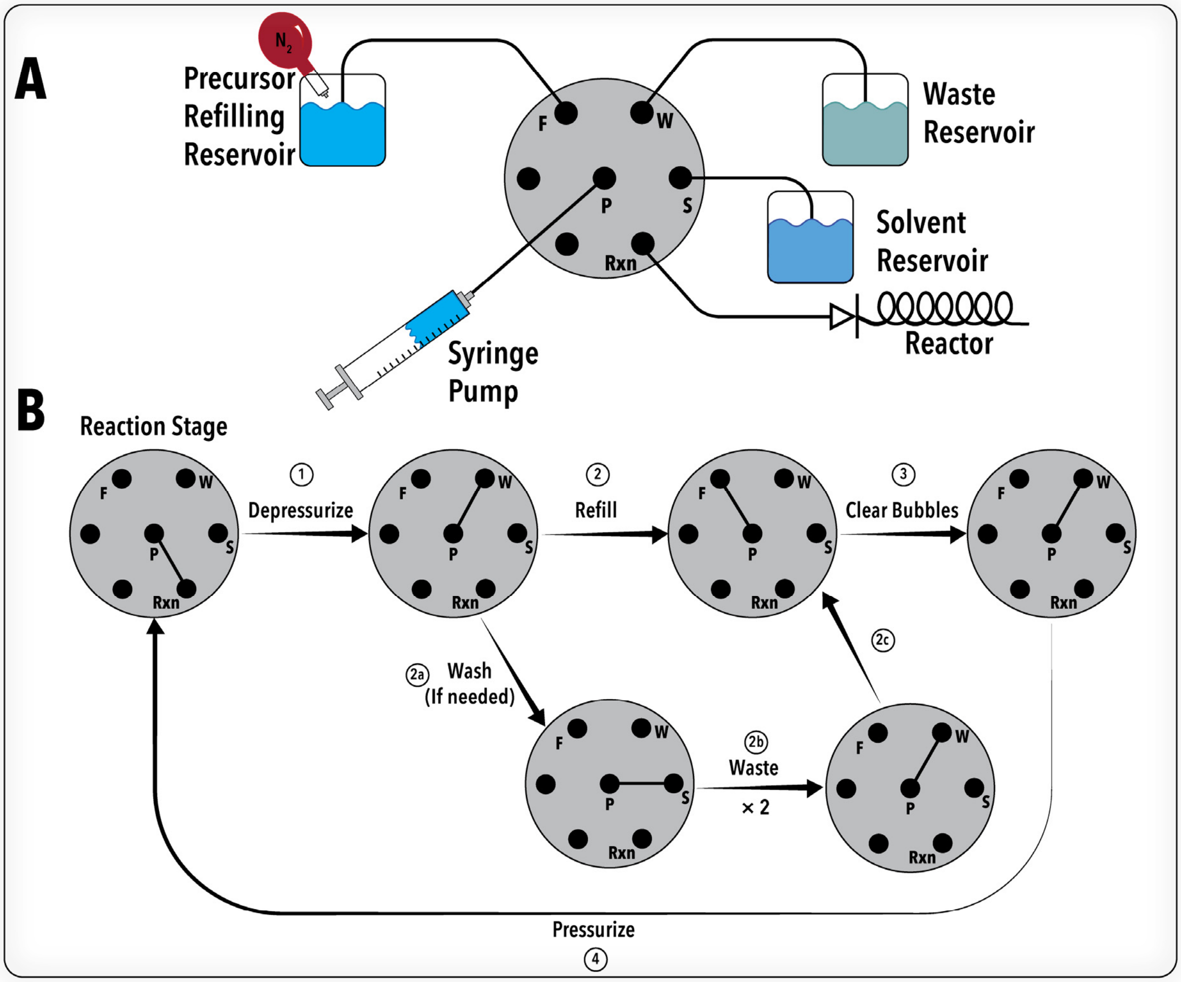

Figure 2. A) Schematic of Fast-Cat's automated precursor refilling module. B) Steps of the automatic precursor refilling process.

Each precursor syringe pump (P) is equipped with an automated refill module consisting of a computer-controlled selector valve connected to the filling reservoir (F), the waste collection vessel (W), the wash solvent (S), and the main reactor line (Rxn), illustrated in **Fig. 2**. Between reaction conditions, the Fast-Cat process automation module (in LabVIEW) checks to see if the current liquid fill of each syringe pump is sufficient to conduct the upcoming reaction condition. If the liquid level is not sufficient in any of the precursor syringes, it automatically triggers a refill of all precursors. The cycle starts with switching the syringe from the reaction line to the waste collection vessel to allow depressurization (step 1). The pump is then pointed at the precursor reservoir and withdraws enough material to overfill the syringe by ~2 mL (step 2). The excess material is then sent to the waste collection to clear any bubbles that may have been introduced to the syringe (step 3). And, then the syringe is reconnected to the reaction line and repressurized to the flow reactor pressure (step 4). If the syringe is to be refilled with a different precursor solution, the syringe is filled with solvent (step 2a) and sent to waste collection (step 2b) two times, before refilling from the new precursor material (step 2c).

4. GC Auto Injection Calibration

To ensure the accuracy of injection samples and prevent sample carryover, thorough washing of the sampling loop in the auto-injection setup is crucial after each injection. In order to assess the efficacy of the washing system, a series of five known samples with incrementally increasing concentrations were injected with the module, while monitoring the GC response. The sampling loop was washed five times with 10 µL of acetone between samples as a part of the central process control module of Fast-Cat (LabVIEW). The linear GC response with respect to increasing concentration in the samples (**Fig. 3**) illustrates the sufficiency of the washing method between the samples to avoid carryover of sample constituents in the sample loop between injections.

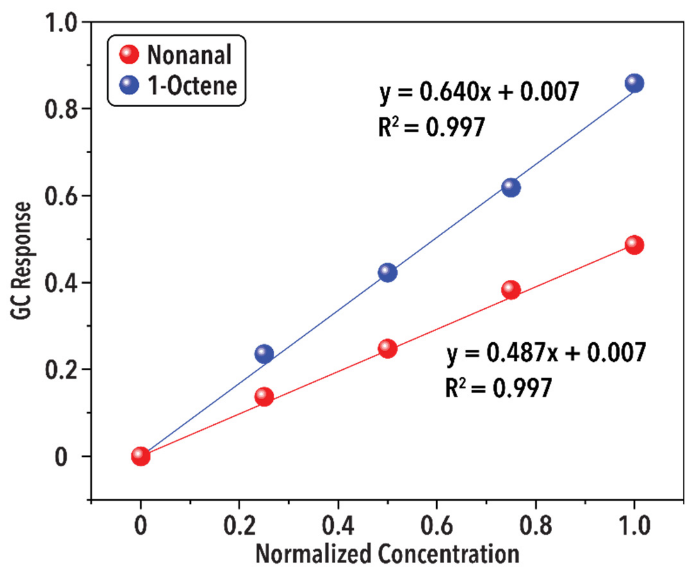

Figure 3. *GC Response over changing concentration in the 1-Octene and nonanal samples in auto-injection, to check efficiency of washing (Five washing cycles between the samples).*

5. In-Flow Mixing

To investigate the effectiveness of liquid-liquid mixing within the flow reactor, the GC responses of manual samples taken at different concentrations were compared with the samples obtained at the outlet of the flow reactor. This experiment is performed to ensure that the liquid precursors flowing from the syringes are adequately mixed inside the flow reactor. **Fig. 4** illustrates the agreement between the response of the in-flow mixed *vs*. the manually well-mixed samples.

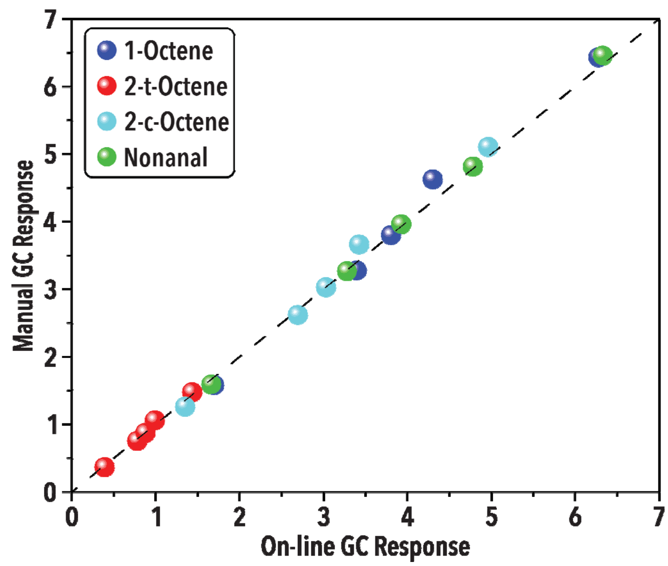

Figure 4. *Comparison between the GC responses of the in-flow mixed vs. the manually prepared samples of 1-octene, 2-trans-octene, 2-cis-octene, and nonanal.*

6. Batch *vs*. Flow Benchmarking

Benchmarking the flow reactor performance of Fast-Cat *vs*. a batch reactor (autoclave) is critical to ensure Fast-Cat's knowledge scalability and industrial relevance. Achieving a benchmark comparison between flow reactor performance and batch operation is challenging, especially at lower reaction temperatures due to the low hydroformylation rates involved. In such cases, isomerization reactions tend to dominate and outperform the hydroformylation process. The difference in the heat and mass transfer rates of a gas-liquid segmented flow configuration in a continuous flow reactor *vs*. a batch reactor can result in discrepancies in the total aldehyde yield and regioselectivities for the same reaction condition. An ideal flow reactor would be able to replicate the batch reactor performance at a shorter reaction time to enable higher reaction exploration throughput per day. Thus, we performed a series of flow reactor geometry (inner diameter), length (volume), and heating mechanism (oil bath *vs*. a heating block) screening using ligan L1 to obtain an ideal flow reactor configuration replicating a batch reactor performance.

Increasing the reaction time can potentially enhance aldehyde yields for slower hydroformylation rates. However, longer reaction times result in lower flow rates of both gas and liquid, consequently hampering the overall mass transfer rates and subsequently decreasing the aldehyde yields. On the other hand, decreasing the inner diameter of the flow reactor tube (from 1/8" to 0.03") proved beneficial in increasing aldehyde yields. This improvement was attributed to the increased flow velocities of gas and liquid segments, enhancing the mass transfer rates. Further optimization was pursued by reducing the inside diameter of the flow reactor tube to 0.02" in. However, this change required a longer flow reactor. In order to extend the residence times of the reactions without impeding the mass transfer rates, increasing the flow reactor volume was deemed necessary (6 ml). Using an oil bath resulted in a lower total aldehyde yield than a heating block. The observed disparity can be attributed to the challenges in achieving uniform heating of the extended flow reactor tube length within an oil bath *vs*. a heating block. The optimal flow reactor configuration was found to be a 6-ml reactor with stainless-steel tubing (1/16" OD ×0.03" ID) placed inside a heating block heated with computer-controlled cartridge heaters. The benchmarking of the optimal flow reactor configuration *vs*. a batch reactor for ligand L1 is presented in **Fig. 5**.

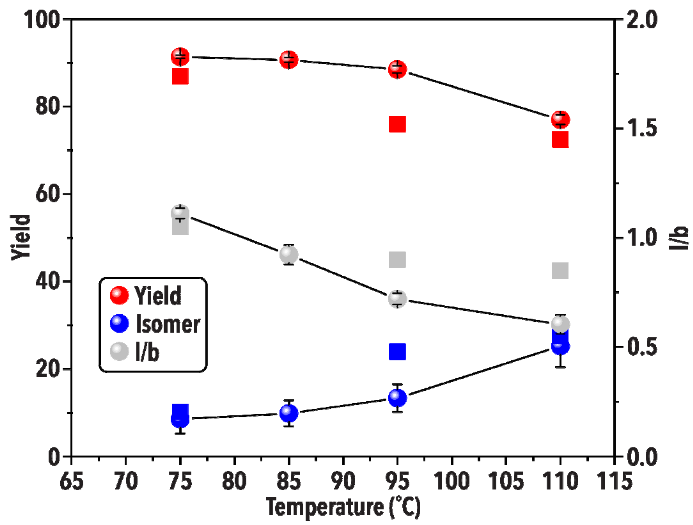

Figure 5. *Benchmarking the optimal flow reactor configuration of Fast-Cat (circles, mean of 3 replicates +/- standard deviation) vs. a batch reactor (squares) with ligand L1. Flow Reaction Conditions: 200 Psig, 1:1 CO:H_2, 10:1 L1:Rh, 0.25 mM Rh, 0.5 M 1-octene, 60 min t_R, 3:1 gas:liquid. Batch Reaction Conditions: 200 Psig, 1:1 CO:H_2, 10:1 L1:Rh, 0.25 mM Rh, 0.5M 1-octene, 120 min t_R.*

7. Auto-Injection Replication

To evaluate the consistency of the custom-developed auto-injection module of Fast-Cat after reaching steady-state operation, the hydroformylation reaction was repeated for multiple residence times. Automated online sampling and washing were conducted using the auto-injection module after each residence time and compared with offline GC injection. **Fig. 6** show the auto-injection replication results for 1-octene isomerization yield, total aldehyde yield, and aldehyde regioselectivity (*l*/*b*) with ligand L1, using the reaction conditions reported in **Table 2**. The results shown in **Fig. 6** demonstrate the accuracy and reliability of the developed auto-injection and washing mechanism of Fast-Cat. The yield of internal octene isomers is 17±1% (6% standard deviation). The yield of aldehydes is 18.5±0.7% (4% standard deviation). The selectivity is 2.56±0.04 l/b (2% standard deviation)

Table 2. *Operating conditions for evaluating the consistency and stability of the auto-injection sampling module of Fast-Cat.*

Temperature (°C)	**Pressure (psig)**	**Ligand (L1):Rh**	**CO:H_2**	**[1-Octene]**	**[Rh]**	**G:L ratio**	**Residence Time (min)**
110	200	10:1	1:1	0.5 M	0.25 mM	2:1	11

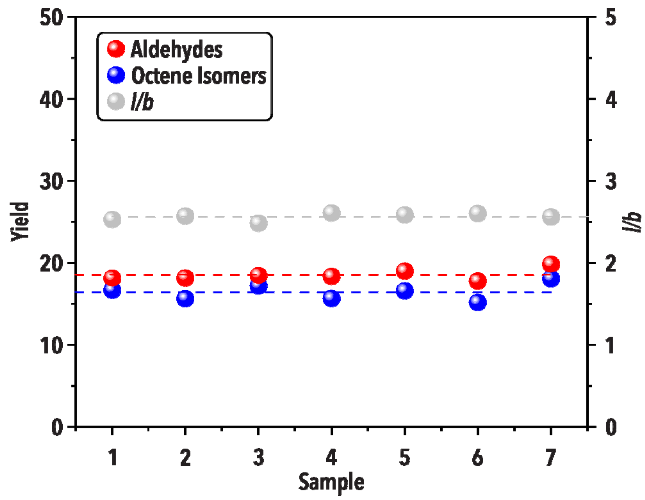

Figure 6. *Total aldehyde yield (%), Octene isomerization yield (%), and Selectivity (l/b) at the reaction conditions in* ***Table S4*** *repeated for number of residence times.*

8. Temperature and Pressure System Response

When performing closed-loop autonomous experimentation, it is common to vary the operating reaction conditions, including reaction temperature and pressure, to explore their effects on the homogeneous catalytic reaction of interest. In order to obtain reliable information after each automatically conducted in-flow hydroformylation reaction, it is important to quantify the characteristic response time needed for each process parameter of Fast-Cat to reach steady-state operation. **Fig. 7** presents the response times for flow reactor temperature changes during extreme heating and cooling processes (from the highest to the lowest reaction temperatures), as well as extreme pressure change during pressure increase (from the lowest to the highest pressure). It is observed that the pressure decrease occurs almost instantaneously upon automatically adjusting the backpressure regulator setpoint. These response times, shown in **Fig. 7**, provide valuable insights into the time taken for the Fast-Cat hardware to achieve stable operating conditions and enable the system to achieve steady-state for reliable and consistent online reaction sampling. Heating the reactor from the minimum to the maximum temperature (75-110°C) reaches steady-state after 10 min, while the inverse cooling takes 25 min due to the insulation **Fig. 7A**. The pressurization of the reactor from the minimum to maximum set point (150-300 psig) takes 6 min using the gas MFCs set to 50% or 1.0 mLn/min each for a total flow of 2.0 mLn/min **Fig. 7B**.

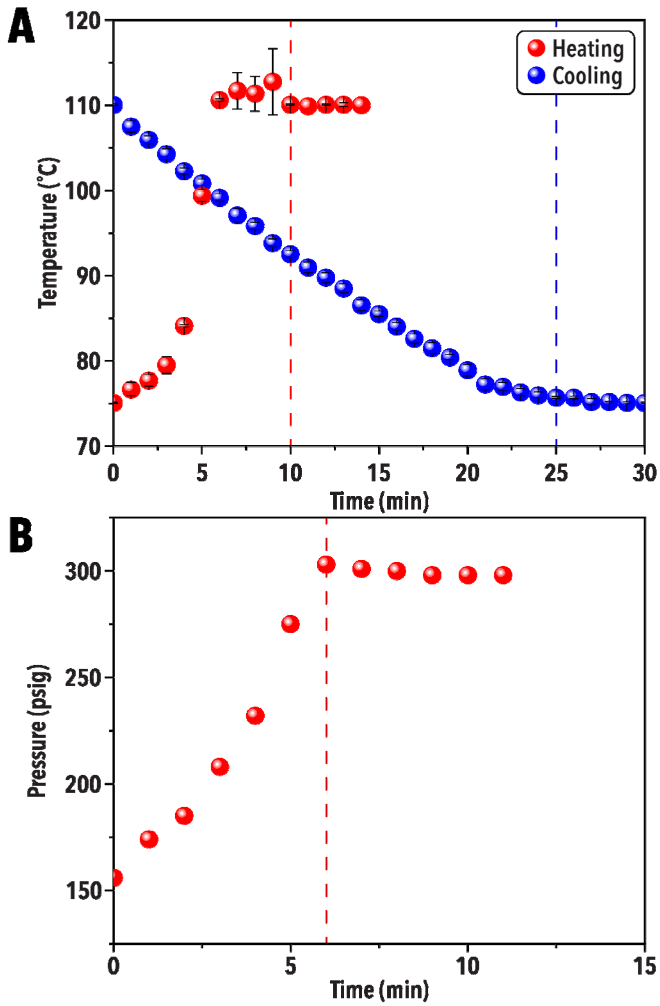

Figure 7. *A) Heating and cooling response for Fast-Cat's flow reactor between 75°C and 110°C (mean over 3 experiments +/- standard deviation). B) Pressurization response of Fast-Cat's flow reactor from 150 psig to 300 psig.*

9. Benchmarking and Validation Conditions.

Table 3. Detailed reaction conditions for results shown in ***Extended Data Figure 2.***

Panel	*Pressure (psig)*	*Temp. °C*	*CO: H_2*	*L1: Rh*	*[Rh] (mM)*	*[1-octene] (M)*	*Gas:liquid*	t_R (min)
1A	200	110	1:1	10:1	0.25	0.5	2:1	45 min
1A	200	85	1:3	10:1	0.25	0.5	2:1	45 min
1B	200	110	1:1	10:1	0.25	0.5	3:1	30 min
1C	200	110	1:1	10:1	0.25	0.5	2:1	11 min
1D 1,7	200	110	1:1	10:1	0.25	0.5	2:1	20 min
1D 2	200	75	1:1	10:1	0.25	0.5	2:1	30 min
1D 3	200	95	2:1	10:1	0.25	0.5	2:1	20 min
1D 4	200	85	1:2	10:1	0.25	0.5	3:1	30 min
1D 5	200	120	1:1	10:1	0.25	0.5	2:1	15 min
1D 6	200	110	1:1	10:1	0.25	0.5	1:1	15 min

10. Model Parameters and Normalization.

Table 4. Summary of model input normalization for converting equipment settings and flow rates to non-dimensionalized parameters between 0 and 1.

	X_1 CO	X_2 H_2	X_3 Pressure	X_4 Temp.	X_5 Dilution	X_6 L:Rh	X_7 Olefin:Rh
Min	0.1 mLn/min	0.1 mLn/min	150 psig	75 °C	0.1	0.2	0.2
Max	1.9 mLn/min	1.9 mLn/min	300 psig	110 °C	0.4	0.8	0.8
Formula	$\frac{X-0.1}{1.9-0.1}$	$\frac{X-0.1}{1.9-0.1}$	$\frac{X-150}{300-150}$	$\frac{X-75}{110-75}$	$\frac{X-0.1}{0.4-0.1}$	$\frac{X-0.2}{0.8-0.2}$	$\frac{X-0.2}{0.8-0.2}$
Def.	MFC Flowrate	MFC Flowrate	Total Pressure	Reactor Temperature	Solvent Stream Fraction	Ligand Stream Fraction	Olefin Stream Fraction
					$\frac{\dot{V}_{Solvent}}{\dot{V}_{liquid}}$	$\frac{\dot{V}_{Ligand}}{\dot{V}_{Ligand}+\dot{V}_{Rh}}$	$\frac{\dot{V}_{Olefin}}{\dot{V}_{Olefin}+\dot{V}_{Rh}}$

Model parameters for Fast-Cat were selected as normalized hardware setpoints to simplify incorporating hardware constraints into the selection of new experimental conditions. The other option for defining the gas phase (total flow rate, pressure, and CO:H2 ratio) results in physically inaccessible regions of parameter space. For example, at a desired gas flowrate of 3.8 mLn/min, it is only possible to have a 1:1 CO:H_2 ratio as both MFCs would be operating at their maximum value, whereas a flowrate of 2 mLn/min could allow for ratios of 1:19 to 19:1 (0.1 mLn/min CO and 1.9 mLn/min H_2 and *vice versa*). When the flowrate drops to 0.2 mLn/min, only a 1:1 syngas ratio becomes possible as the MFCs are operating at their minimum value. Given the set of parameters, gas and liquid stream compositions, flowrates, temperature, pressure, and residence time are fully defined for a constant flow reactor volume and gas:liquid volumetric ratio of 3:1 in the pressurized reactor. A consequence of the selected parameters is that X_1-X_3 contribute to the total residence time in the flow reactor either as a greater total flow rate or expansion/compression of the gas phase within the flow reactor.

11. qNEHVI Ligand Screening Data.

Table 5. *Summary of experimental conditions autonomously selected and tested by Fast-Cat for autonomous Pareto-front mapping of ligand L1. Init: initialization run; N: Normal aldehyde campaign; I: Iso aldehyde campaign.*

X_1 CO	X_2 H_2	X_3 Pressure	X_4 Temp.	X_5 Dilution	X_6 L:Rh	X_7 Olefin:Rh	Y Yield	S_N Selectivity	Opt.
0.309	0.309	0.347	0.000	0.500	0.500	0.500	0.914	0.528	Init.
0.674	0.674	0.347	1.000	0.500	0.500	0.500	0.684	0.416	Init.
0.309	0.309	0.347	1.000	0.500	0.500	0.500	0.770	0.376	Init.
0.431	0.917	0.347	0.286	0.500	0.500	0.500	0.687	0.571	Init.
1.000	0.806	0.966	0.252	1.000	0.096	0.428	0.784	0.626	N
1.000	0.000	0.359	0.723	0.000	0.343	1.000	0.081	0.709	N
0.793	1.000	0.819	0.000	0.395	0.992	0.980	0.254	0.673	N
0.000	1.000	1.000	0.000	0.739	0.428	1.000	0.446	0.823	N
0.504	0.000	0.000	0.119	0.876	1.000	0.869	0.101	0.688	N
0.999	0.882	0.567	0.481	0.984	1.000	0.320	0.721	0.662	N
0.000	1.000	0.897	0.990	0.282	1.000	0.591	0.747	0.603	N
0.263	0.874	0.991	1.000	1.000	0.845	0.597	0.976	0.473	N
0.000	1.000	0.387	0.000	1.000	1.000	0.176	0.603	0.813	N
1.000	0.901	1.000	0.000	0.388	1.000	0.000	0.890	0.652	N
0.748	1.000	1.000	0.825	1.000	0.879	0.000	0.915	0.526	N
0.093	0.539	0.000	0.000	0.861	1.000	1.000	0.541	0.770	N
0.911	1.000	0.813	0.000	0.944	0.500	0.949	0.218	0.678	N
0.802	0.719	0.000	0.000	1.000	0.595	1.000	0.038	0.701	N
0.000	1.000	1.000	0.000	0.700	0.838	0.921	0.530	0.838	N
0.090	1.000	0.668	0.000	0.287	1.000	1.000	0.611	0.789	N
0.000	1.000	0.000	0.458	0.769	0.947	1.000	0.178	0.844	N
0.000	1.000	1.000	0.000	1.000	1.000	0.900	0.582	0.851	N
0.936	1.000	1.000	0.773	0.372	1.000	1.000	0.788	0.563	N

1.000	1.000	0.000	0.070	0.000	0.979	0.101	0.340	0.717	N
0.202	1.000	0.123	1.000	0.918	1.000	0.533	0.432	0.593	N
1.000	0.707	1.000	0.741	1.000	0.822	0.000	1.000	0.540	N
0.411	1.000	0.958	0.961	1.000	1.000	0.773	0.851	0.518	N
0.000	1.000	1.000	0.825	0.000	1.000	1.000	0.561	0.751	N
1.000	1.000	0.603	0.859	0.542	0.000	0.038	0.776	0.506	N
0.070	1.000	0.337	0.000	0.392	1.000	1.000	0.525	0.832	N
0.000	1.000	0.617	0.075	0.113	0.719	0.977	0.362	0.857	N
0.000	1.000	0.979	0.000	1.000	1.000	0.480	0.682	0.802	N
1.000	0.000	0.466	0.338	0.922	1.000	1.000	0.105	0.701	N
1.000	0.855	0.000	0.758	1.000	0.827	1.000	0.409	0.647	N
0.439	0.542	1.000	0.638	0.844	0.217	0.000	1.000	0.504	I
0.533	0.790	1.000	0.528	1.000	0.218	0.000	0.988	0.514	I
0.940	0.629	1.000	0.509	1.000	1.000	0.148	0.834	0.591	I
0.372	0.586	1.000	0.941	0.392	0.000	0.000	1.000	0.477	I
0.228	1.000	0.712	0.683	0.532	1.000	0.112	0.880	0.578	I
1.000	0.971	1.000	1.000	0.269	0.847	0.000	0.934	0.506	I
0.439	1.000	0.529	1.000	1.000	1.000	0.337	0.664	0.548	I
0.550	1.000	0.763	1.000	1.000	0.757	0.221	0.945	0.483	I
0.360	1.000	1.000	1.000	1.000	0.000	0.132	0.920	0.470	I
1.000	0.000	1.000	0.356	1.000	0.377	0.000	0.820	0.555	I
0.815	0.631	1.000	1.000	0.509	0.000	0.000	0.893	0.467	I
0.699	1.000	1.000	0.922	1.000	0.841	0.000	0.962	0.491	I
0.933	0.932	1.000	0.661	0.041	0.558	0.000	0.994	0.517	I
0.402	0.597	0.977	1.000	0.662	0.282	0.000	0.995	0.463	I
0.455	0.951	1.000	1.000	0.975	0.384	0.000	1.000	0.446	I
1.000	1.000	0.000	0.000	0.000	1.000	1.000	0.049	0.712	I
1.000	0.841	0.066	0.003	0.797	1.000	0.995	0.046	0.702	I
0.000	1.000	1.000	1.000	0.905	0.365	0.586	0.713	0.551	I
0.922	0.823	0.429	0.852	1.000	1.000	1.000	0.685	0.561	I
0.074	0.997	1.000	0.000	1.000	0.809	0.152	0.897	0.669	I
1.000	1.000	1.000	1.000	0.956	0.146	0.074	0.895	0.469	I
1.000	1.000	1.000	1.000	0.467	0.000	0.000	0.894	0.480	I
0.303	0.712	1.000	0.000	0.637	0.842	0.742	0.946	0.613	I
0.379	1.000	1.000	0.147	0.000	1.000	0.000	0.898	0.642	I
0.634	1.000	0.886	1.000	1.000	1.000	1.000	0.781	0.500	I
0.497	0.493	0.924	0.135	1.000	0.701	0.922	0.904	0.619	I
0.635	0.336	0.890	0.172	1.000	1.000	1.000	0.639	0.678	I
0.178	0.814	1.000	0.777	1.000	1.000	0.232	0.888	0.555	I
0.000	1.000	1.000	0.052	1.000	0.631	0.789	0.650	0.808	I
0.685	1.000	0.875	0.000	1.000	0.234	0.497	0.889	0.635	I
1.000	0.822	1.000	0.000	1.000	1.000	1.000	0.090	0.663	I

Table 6. *Summary of experimental conditions autonomously selected and tested by Fast-Cat for autonomous Pareto-front mapping of ligand L2. Init: initialization run; N: Normal aldehyde campaign; I: Iso aldehyde campaign.*

X_1 CO	X_2 H_2	X_3 Pressure	X_4 Temp.	X_5 Dilution	X_6 L:Rh	X_7 Olefin:Rh	Y Yield	S_N Selectivity	Opt.
0.222	0.778	1.000	0.000	1.000	1.000	0.000	0.873	0.741	Init.
0.778	0.222	1.000	1.000	0.000	0.000	1.000	0.261	0.459	Init.
0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.941	0.612	Init.
1.000	1.000	0.000	0.750	0.750	0.250	0.250	0.646	0.640	Init.
1.000	1.000	1.000	0.223	1.000	0.957	0.855	0.839	0.687	N
0.804	1.000	1.000	0.990	1.000	0.697	0.131	0.938	0.607	N
0.623	0.658	1.000	1.000	0.471	0.866	0.681	0.905	0.607	N
0.881	0.699	0.808	0.392	0.940	1.000	1.000	0.840	0.705	N
0.334	0.680	1.000	0.019	1.000	1.000	0.656	0.878	0.712	N
1.000	0.819	0.938	0.000	1.000	1.000	0.130	0.656	0.685	N
0.178	1.000	0.026	0.203	1.000	1.000	0.919	0.555	0.859	N
0.699	0.919	0.679	0.941	0.985	0.727	0.700	0.832	0.605	N
1.000	0.819	0.752	0.804	0.771	0.987	1.000	0.772	0.638	N
0.406	1.000	0.000	0.900	0.215	0.875	1.000	0.618	0.718	N
0.884	0.564	0.575	0.749	1.000	0.539	1.000	0.691	0.590	N
0.956	0.295	0.317	0.884	0.626	0.818	1.000	0.556	0.645	N
1.000	0.730	0.292	0.000	0.000	1.000	0.000	0.400	0.626	N
0.283	1.000	0.932	0.765	0.310	0.326	1.000	0.877	0.600	N
0.734	0.824	1.000	0.490	0.987	0.865	0.588	0.893	0.683	N
0.672	0.139	0.125	0.638	1.000	1.000	1.000	0.565	0.751	N
1.000	0.924	0.000	0.806	0.744	1.000	0.752	0.605	0.787	N
0.870	0.995	0.937	0.772	0.932	1.000	0.765	0.847	0.702	N
1.000	0.984	0.491	0.058	0.408	1.000	0.654	0.410	0.667	N
0.012	0.934	0.000	0.088	0.855	0.464	1.000	0.247	0.915	N
0.515	0.254	0.257	1.000	0.439	0.921	1.000	0.653	0.618	N
0.937	1.000	0.200	0.625	1.000	1.000	1.000	0.606	0.775	N
0.000	0.858	0.754	1.000	0.658	1.000	0.960	0.675	0.829	N
0.532	0.781	1.000	0.244	0.581	1.000	0.000	0.894	0.724	I
1.000	1.000	1.000	0.934	0.419	1.000	0.791	0.834	0.674	I
0.625	0.800	1.000	0.735	0.380	0.667	1.000	0.948	0.553	I
0.178	1.000	0.376	0.238	0.618	1.000	1.000	0.665	0.781	I
0.700	0.916	1.000	0.144	1.000	0.293	1.000	0.708	0.537	I
1.000	1.000	0.182	0.915	1.000	1.000	0.010	0.650	0.778	I
0.888	0.648	1.000	0.810	1.000	0.717	0.748	0.896	0.556	I
1.000	0.556	0.401	0.677	1.000	1.000	1.000	0.584	0.699	I
0.751	1.000	0.264	0.658	1.000	1.000	1.000	0.494	0.612	I
1.000	0.473	1.000	0.759	0.976	0.066	0.780	0.245	0.454	I
0.776	1.000	1.000	1.000	1.000	0.501	0.720	0.292	0.454	I

0.412	1.000	1.000	1.000	0.776	0.618	0.871	0.881	0.531	I
1.000	0.000	0.736	0.500	0.251	0.888	0.406	0.587	0.640	I
1.000	0.891	0.951	0.502	0.881	1.000	0.710	0.780	0.674	I
0.616	0.580	1.000	0.581	0.794	1.000	0.849	0.867	0.607	I
0.000	1.000	1.000	0.174	0.000	0.739	0.350	0.880	0.782	I
0.840	0.508	0.868	0.106	0.875	1.000	0.427	0.809	0.690	I
0.692	0.506	1.000	0.116	0.994	0.837	0.487	0.905	0.679	I
1.000	1.000	1.000	0.379	0.851	0.817	0.173	0.858	0.700	I
0.504	1.000	1.000	0.731	0.362	1.000	0.000	0.813	0.731	I
0.491	0.662	0.976	1.000	0.270	0.892	0.655	0.904	0.634	I

Table 7. *Summary of experimental conditions autonomously selected and tested by Fast-Cat for autonomous Pareto-front mapping of ligand L3. Init: initialization run; N: Normal aldehyde campaign; I: Iso aldehyde campaign.*

X_1 CO	X_2 H_2	X_3 Pressure	X_4 Temp.	X_5 Dilution	X_6 L:Rh	X_7 Olefin:Rh	Y Yield	S_N Selectivity	Opt.
0.222	0.778	1.000	0.000	1.000	1.000	0.000	0.924	0.749	Init.
0.778	0.222	1.000	1.000	0.000	0.000	1.000	0.444	0.535	Init.
0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.844	0.644	Init.
1.000	1.000	0.000	0.750	0.750	0.250	0.250	0.492	0.704	Init.
0.586	0.000	0.063	0.452	1.000	1.000	0.778	0.407	0.783	N
0.658	0.478	0.000	0.743	1.000	1.000	1.000	0.479	0.770	N
0.883	0.778	0.959	1.000	1.000	0.554	0.546	0.848	0.590	N
1.000	0.857	0.924	0.926	0.868	0.684	0.818	0.781	0.618	N
0.493	1.000	0.550	0.943	0.876	1.000	0.384	0.683	0.731	N
1.000	0.000	1.000	1.000	1.000	1.000	0.000	0.609	0.623	N
0.429	1.000	1.000	0.010	1.000	0.841	0.748	0.870	0.737	N
0.411	1.000	1.000	0.530	1.000	0.993	0.854	0.823	0.712	N
1.000	1.000	1.000	0.820	0.000	0.467	0.085	0.882	0.625	N
0.390	1.000	1.000	0.000	0.877	1.000	0.712	0.843	0.760	N
0.676	0.945	1.000	0.325	0.820	1.000	1.000	0.858	0.739	N
0.000	1.000	1.000	0.000	1.000	0.536	0.024	0.808	0.802	N
0.909	1.000	1.000	0.835	1.000	0.883	0.764	0.790	0.658	N
0.906	0.981	1.000	0.655	1.000	0.886	0.747	0.812	0.686	N
0.560	1.000	1.000	0.100	0.971	0.588	0.837	0.908	0.704	N
0.000	1.000	0.430	0.995	0.000	0.563	1.000	0.174	0.882	N
0.518	1.000	1.000	0.652	0.896	1.000	0.542	0.862	0.709	N
0.002	1.000	1.000	0.468	1.000	0.986	0.000	0.845	0.828	N
0.482	0.871	0.057	0.508	1.000	0.769	0.000	0.636	0.784	N
1.000	0.537	0.048	1.000	1.000	1.000	0.209	0.432	0.760	N
0.449	1.000	0.720	0.847	1.000	1.000	0.283	0.693	0.717	I
0.607	1.000	0.739	0.747	1.000	1.000	0.648	0.797	0.735	I
0.882	0.373	1.000	0.650	1.000	0.082	0.642	0.574	0.656	I
0.397	0.754	0.203	0.000	1.000	1.000	0.566	0.664	0.800	I
0.685	0.647	0.767	0.899	0.000	0.330	0.871	0.707	0.593	I
1.000	0.963	0.579	0.838	0.671	0.720	1.000	0.601	0.663	I
0.000	0.656	0.393	0.605	0.438	0.766	1.000	0.334	0.907	I
0.757	0.729	1.000	0.543	0.555	1.000	0.768	0.869	0.691	I
0.475	1.000	0.967	0.000	0.465	1.000	0.078	0.873	0.740	I
0.908	0.782	1.000	0.852	0.623	1.000	0.034	0.831	0.693	I
0.629	0.800	1.000	0.554	0.862	1.000	0.685	0.857	0.692	I
0.026	1.000	0.911	0.000	0.905	0.485	0.825	0.750	0.900	I
0.902	0.000	0.425	0.758	1.000	0.932	0.882	0.385	0.745	I
0.445	1.000	1.000	0.985	1.000	0.912	0.688	0.404	0.577	I

0.018	1.000	1.000	0.582	1.000	0.404	0.001	0.483	0.656	I
0.992	1.000	1.000	0.517	0.552	1.000	0.640	0.882	0.695	I
1.000	0.949	0.794	0.633	0.130	0.000	0.000	0.828	0.587	I
1.000	0.571	0.939	0.000	0.102	0.274	0.131	0.898	0.669	I
0.857	0.868	0.848	0.068	0.479	1.000	0.000	0.889	0.719	I

Table 8. Summary of experimental conditions autonomously selected and tested by Fast-Cat for autonomous Pareto-front mapping of ligand L4. Init: initialization run; N: Normal aldehyde campaign; I: Iso aldehyde campaign.

X_1 CO	X_2 H_2	X_3 Pressure	X_4 Temp.	X_5 Dilution	X_6 L:Rh	X_7 Olefin:Rh	Y Yield	S_N Selectivity	Opt.
0.222	0.778	1.000	0.000	1.000	1.000	0.000	0.994	0.741	Init.
0.778	0.222	1.000	1.000	0.000	0.000	1.000	0.210	0.453	Init.
0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.976	0.728	Init.
1.000	1.000	0.000	0.750	0.750	0.250	0.250	0.874	0.721	Init.
0.000	1.000	0.082	0.646	1.000	0.549	0.719	0.404	0.805	N
0.613	0.645	0.500	0.573	0.246	0.931	0.863	0.944	0.739	N
0.981	1.000	0.303	0.520	1.000	0.969	0.195	0.886	0.743	N
0.758	1.000	1.000	0.284	0.883	1.000	0.455	0.926	0.733	N
0.439	0.458	0.516	0.959	1.000	1.000	1.000	0.887	0.729	N
0.473	1.000	0.492	0.440	1.000	0.898	1.000	0.827	0.732	N
1.000	0.709	0.697	0.590	0.523	1.000	0.785	0.796	0.741	N
0.661	1.000	0.917	1.000	0.492	0.953	0.652	0.951	0.724	N
0.906	1.000	1.000	0.732	1.000	1.000	1.000	0.801	0.730	N
0.184	1.000	1.000	1.000	0.666	1.000	0.000	1.000	0.741	N
1.000	1.000	0.975	0.681	0.918	0.360	0.774	0.866	0.724	N
1.000	1.000	0.350	0.866	1.000	0.971	0.398	0.956	0.749	N
0.402	1.000	0.321	0.915	0.000	1.000	0.876	0.922	0.752	N
1.000	1.000	0.621	0.000	0.712	1.000	1.000	0.093	0.773	N
1.000	1.000	0.234	0.857	0.675	0.696	1.000	0.632	0.741	N
0.891	0.911	0.208	0.868	0.000	1.000	1.000	0.798	0.742	N
0.144	0.755	0.530	0.096	1.000	1.000	0.221	0.961	0.760	N
0.784	0.570	0.650	1.000	0.680	1.000	0.849	0.939	0.727	N
0.008	1.000	1.000	0.625	0.000	0.502	0.899	0.360	0.794	N
0.481	0.765	0.787	0.913	1.000	1.000	0.530	0.948	0.727	N
0.132	0.662	0.613	0.569	1.000	1.000	0.301	1.000	0.777	N
0.260	1.000	0.867	0.967	1.000	1.000	0.432	0.968	0.760	N
0.906	1.000	0.922	1.000	0.637	1.000	0.444	0.929	0.731	N
0.871	1.000	0.942	1.000	1.000	0.000	0.000	0.751	0.634	N
0.463	1.000	1.000	0.303	0.320	1.000	0.000	1.000	0.741	I
0.767	0.544	0.130	1.000	1.000	1.000	0.225	0.867	0.739	I
1.000	0.000	1.000	0.029	0.963	1.000	0.704	0.050	0.741	I
0.658	1.000	0.638	1.000	1.000	0.811	0.000	0.915	0.706	I
0.430	0.714	0.095	0.888	0.442	0.000	1.000	0.093	0.664	I
0.924	0.417	0.367	0.002	0.562	1.000	0.911	0.089	0.774	I
0.563	1.000	1.000	0.347	1.000	1.000	0.000	1.000	0.734	I
0.976	1.000	0.000	1.000	1.000	0.362	1.000	0.668	0.723	I
0.506	0.801	1.000	0.680	0.586	0.648	0.356	0.970	0.725	I
0.707	1.000	0.774	1.000	0.573	1.000	0.671	0.953	0.727	I

0.000	1.000	0.000	0.243	1.000	1.000	0.241	0.930	0.781	I
1.000	0.870	0.000	1.000	0.455	0.779	0.000	0.884	0.743	I
0.543	1.000	1.000	0.000	0.805	1.000	0.000	0.831	0.701	I
1.000	0.864	0.741	0.759	1.000	0.586	0.499	0.963	0.720	I
0.964	1.000	0.000	0.813	1.000	0.753	0.000	0.926	0.744	I
0.000	1.000	1.000	0.562	0.737	1.000	0.079	0.933	0.759	I
0.381	1.000	0.467	1.000	1.000	1.000	0.106	0.949	0.726	I
0.948	1.000	0.466	0.737	0.747	0.659	0.440	0.961	0.730	I
1.000	0.504	1.000	0.631	0.790	0.570	0.603	0.952	0.715	I
0.771	1.000	0.892	0.718	0.722	0.859	0.000	0.994	0.741	I
0.619	0.755	0.508	1.000	0.312	1.000	0.462	0.951	0.742	I
0.835	1.000	0.578	0.823	0.928	1.000	0.370	0.931	0.721	I
0.200	0.511	0.329	0.361	0.961	0.797	0.000	0.949	0.699	I
0.755	0.996	1.000	0.829	0.750	0.768	0.812	0.965	0.727	I
0.345	1.000	0.872	0.641	1.000	1.000	0.402	0.944	0.738	I

Table 9. Summary of experimental conditions autonomously selected and tested by Fast-Cat for autonomous Pareto-front mapping of ligand L5. Init: initialization run; N: Normal aldehyde campaign; I: Iso aldehyde campaign.

X_1 CO	X_2 H_2	X_3 Pressure	X_4 Temp.	X_5 Dilution	X_6 L:Rh	X_7 Olefin:Rh	Y Yield	S_N Selectivity	Opt.
0.222	0.778	1.000	0.000	1.000	1.000	0.000	0.023	1.000	Init.
0.778	0.222	1.000	1.000	0.000	0.000	1.000	0.107	0.595	Init.
0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.116	1.000	Init.
1.000	1.000	0.000	0.750	0.750	0.250	0.250	0.130	1.000	Init.
0.584	0.518	0.624	0.823	0.207	0.020	0.907	0.177	0.643	N
0.448	0.689	0.406	0.549	0.249	0.016	0.974	0.202	0.797	N
0.842	1.000	0.735	0.601	0.916	0.108	0.301	0.123	0.965	N
0.807	0.934	0.000	0.450	0.905	0.651	0.253	0.032	1.000	N
0.762	0.790	0.455	1.000	0.703	0.112	0.761	0.195	0.978	N
0.296	1.000	0.000	1.000	0.467	0.024	0.079	0.529	0.971	N
0.000	1.000	0.185	0.954	0.000	1.000	0.876	0.253	0.964	N
0.089	1.000	0.209	1.000	0.750	0.070	1.000	0.317	0.957	N
0.871	0.456	0.359	1.000	0.587	0.354	1.000	0.157	0.985	N
0.999	1.000	0.263	1.000	0.783	0.815	0.938	0.075	0.979	N
0.765	1.000	0.438	1.000	1.000	0.070	0.000	0.302	0.980	N
0.429	1.000	0.296	1.000	0.805	0.000	0.365	0.402	0.983	N
0.050	1.000	0.000	1.000	0.439	0.360	0.503	0.393	0.964	N
0.520	1.000	0.497	1.000	0.004	0.602	0.000	0.399	0.981	N
0.099	0.713	0.473	1.000	0.613	1.000	0.399	0.342	0.984	N
0.561	1.000	0.000	0.914	0.193	0.038	0.879	0.237	0.983	N
0.052	1.000	0.236	1.000	0.325	0.000	0.000	0.721	0.980	N
0.000	0.926	0.905	1.000	0.290	0.165	0.471	0.320	0.962	N
0.000	0.862	0.214	1.000	0.968	0.041	0.202	0.372	0.956	N
0.000	1.000	0.000	1.000	0.000	0.606	0.000	0.498	0.970	N
0.000	1.000	0.646	1.000	0.991	0.702	0.688	0.310	0.973	N
0.697	1.000	0.133	1.000	0.413	0.650	0.611	0.202	0.983	N
0.391	1.000	0.000	1.000	1.000	0.167	1.000	0.141	0.982	N
0.155	1.000	0.546	1.000	0.517	0.292	0.000	0.671	0.982	N
0.845	1.000	0.739	1.000	0.581	0.000	0.329	0.316	0.977	N
0.960	0.000	0.592	1.000	1.000	0.362	0.000	0.273	0.975	N
0.245	0.814	0.379	1.000	0.248	0.062	0.000	0.695	0.981	I
0.082	0.946	0.476	0.850	0.150	0.110	0.000	0.788	0.982	I
1.000	0.000	0.000	0.031	0.649	0.000	0.084	0.035	0.976	I
0.000	0.938	1.000	0.730	0.000	1.000	0.244	0.420	0.984	I
0.619	0.818	0.786	0.167	0.923	0.759	0.200	0.033	1.000	I
0.153	0.742	0.630	0.573	0.506	0.093	0.946	0.236	0.984	I
0.000	1.000	1.000	1.000	0.615	0.728	0.313	0.664	0.977	I
1.000	0.000	0.317	1.000	0.000	0.171	0.595	0.305	0.973	I

0.000	1.000	0.653	1.000	0.083	0.000	0.236	0.565	0.954	I
0.219	1.000	1.000	1.000	0.488	0.063	0.000	0.738	0.979	I
0.000	0.716	0.525	1.000	0.034	1.000	0.003	0.633	0.970	I
0.000	0.711	0.102	1.000	0.000	0.000	0.000	0.328	0.966	I
0.072	0.752	0.125	1.000	0.000	0.766	0.000	0.641	0.985	I
0.000	0.876	0.412	1.000	0.654	0.000	0.137	0.503	0.926	I
0.093	0.989	0.776	0.799	0.001	0.221	0.074	0.708	0.981	I
0.119	0.895	0.000	1.000	0.000	0.033	0.054	0.728	0.976	I
0.000	0.880	0.735	1.000	0.688	0.459	0.000	0.672	0.978	I
0.000	1.000	0.279	1.000	0.208	0.522	0.062	0.544	0.970	I
0.000	0.897	0.272	1.000	0.583	0.385	0.000	0.677	0.972	I
0.942	0.519	0.807	0.783	0.000	0.012	0.000	0.473	0.971	I
0.848	0.357	0.017	0.875	0.417	0.764	0.535	0.192	0.986	I
0.534	1.000	1.000	1.000	0.000	0.000	0.000	0.732	0.978	I
0.728	1.000	0.204	0.861	0.000	0.819	0.000	0.253	0.868	I

Table 10. *Summary of experimental conditions autonomously selected and tested by Fast-Cat for autonomous Pareto-front mapping of ligand L6. Init: initialization run; N: Normal aldehyde campaign; I: Iso aldehyde campaign.*

X_1 CO	X_2 H_2	X_3 Pressure	X_4 Temp.	X_5 Dilution	X_6 L:Rh	X_7 Olefin:Rh	Y Yield	S_N Selectivity	Opt.
0.222	0.778	1.000	0.000	1.000	1.000	0.000	0.971	0.592	Init.
0.778	0.222	1.000	1.000	0.000	0.000	1.000	0.304	0.417	Init.
0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.935	0.476	Init.
1.000	1.000	0.000	0.750	0.750	0.250	0.250	0.580	0.454	Init.
0.464	1.000	1.000	0.072	1.000	1.000	0.839	0.898	0.614	N
0.784	1.000	0.769	0.350	1.000	0.596	1.000	0.938	0.525	N
0.969	0.139	0.943	0.000	1.000	0.052	0.256	0.995	0.548	N
1.000	0.432	0.497	0.400	0.828	0.876	1.000	1.000	0.517	N
1.000	0.822	1.000	0.000	0.837	0.982	0.290	0.947	0.610	N
0.016	0.980	0.220	1.000	0.069	1.000	1.000	0.823	0.440	N
0.210	1.000	1.000	0.000	1.000	0.000	0.356	1.000	0.578	N
1.000	0.425	1.000	0.248	1.000	0.082	0.142	0.933	0.548	N
1.000	0.760	1.000	1.000	1.000	0.868	0.876	0.935	0.420	N
0.674	1.000	0.841	0.060	1.000	0.901	0.431	1.000	0.603	N
0.972	0.792	1.000	0.863	0.880	1.000	0.221	1.000	0.484	N
1.000	0.762	0.467	0.203	0.164	0.993	0.941	0.348	0.540	N
0.982	0.754	1.000	0.000	0.965	0.000	0.567	0.928	0.562	N
1.000	1.000	0.081	0.273	1.000	0.000	0.838	0.566	0.583	N
1.000	0.082	1.000	0.151	0.876	1.000	0.371	0.962	0.594	N
0.550	0.115	0.305	0.186	1.000	0.947	1.000	0.805	0.576	N
0.642	0.907	0.747	0.573	0.752	1.000	1.000	0.897	0.484	N
0.200	1.000	0.731	1.000	0.344	0.000	0.235	0.959	0.416	N
0.000	1.000	1.000	0.686	0.791	0.259	0.000	1.000	0.490	N
0.187	0.748	1.000	0.420	1.000	1.000	0.882	0.976	0.509	N
0.142	1.000	0.123	0.003	1.000	1.000	0.585	0.971	0.647	N
1.000	0.067	0.667	0.216	1.000	0.043	0.980	0.538	0.619	N
1.000	1.000	1.000	0.860	0.910	0.339	0.000	0.924	0.434	N
0.420	1.000	1.000	1.000	0.816	0.413	0.211	1.000	0.387	N
1.000	0.547	0.731	0.000	1.000	0.788	0.139	0.964	0.614	N
0.961	0.479	0.333	0.179	1.000	1.000	0.122	0.444	0.613	N
0.000	0.792	0.721	0.000	1.000	0.339	0.205	0.959	0.558	N
1.000	0.113	0.817	0.000	1.000	0.959	0.956	0.955	0.641	N
0.919	0.409	1.000	0.000	1.000	0.997	0.299	1.000	0.608	N
0.572	0.000	0.000	1.000	1.000	0.952	0.000	1.000	0.545	I
1.000	0.560	1.000	0.165	1.000	1.000	0.677	0.916	0.616	I
0.709	0.389	0.957	0.408	1.000	0.384	1.000	0.821	0.492	I
1.000	0.910	0.000	1.000	0.000	0.523	0.686	0.667	0.434	I
0.065	1.000	0.196	0.000	1.000	1.000	0.806	1.000	0.544	I

0.000	1.000	1.000	0.000	1.000	1.000	0.000	1.000	0.602	I
1.000	1.000	0.821	0.949	0.865	0.860	0.740	1.000	0.383	I
0.714	0.910	1.000	0.000	0.526	0.672	0.000	1.000	0.621	I
0.531	0.961	0.530	0.678	0.937	0.840	0.883	0.972	0.497	I
0.704	1.000	1.000	0.000	0.262	0.487	0.000	1.000	0.575	I
1.000	1.000	0.864	0.000	0.195	0.339	0.000	1.000	0.567	I
0.251	1.000	0.836	0.000	0.468	1.000	1.000	0.689	0.557	I
0.000	1.000	1.000	0.675	0.616	0.833	1.000	0.669	0.589	I
0.867	0.798	0.396	0.000	0.000	0.228	0.424	0.997	0.551	I
1.000	0.000	0.065	0.211	0.934	1.000	0.000	0.930	0.609	I
0.570	0.729	1.000	0.000	1.000	0.649	0.238	1.000	0.574	I
1.000	0.274	1.000	0.742	1.000	0.062	0.000	1.000	0.453	I
1.000	0.819	0.945	0.305	0.769	1.000	0.000	1.000	0.590	I
1.000	1.000	0.225	0.229	0.000	0.527	0.349	1.000	0.481	I
1.000	1.000	1.000	0.000	0.000	0.707	0.895	1.000	0.581	I
1.000	0.908	0.607	0.537	1.000	0.973	0.780	1.000	0.460	I
1.000	1.000	1.000	1.000	0.287	1.000	0.000	1.000	0.527	I
0.157	0.978	0.000	1.000	0.056	1.000	0.628	0.973	0.383	I
1.000	1.000	0.000	0.265	1.000	0.000	0.000	0.915	0.553	I
1.000	1.000	0.960	0.000	1.000	0.000	0.000	0.925	0.485	I

12. Knowledge Scalability.

Table 11. *Fast-Cat's knowledge scalability from the flow to batch reactor. Batch reaction time is 2 hr.*

	CO	**H_2**	***P* (PSIG)**	***T* (°C)**	**[L]**	**[RH]**	**[1-OCT]**	***Y***	**S_N**
ISO FLOW L1	0.92 mL/min	1.81 mL/min	300	110	2.3 mM	0.3mM	0.15M	1.0	0.45
ISO BATCH L1	101 psig	199 psig	300	110	2.3 mM	0.3mM	0.15M	1.0	0.47
NORMAL FLOW L5	0.38 mL/min	1.9 mL/min	232	110	2.4 mM	0.4mM	0.2M	0.67	0.98
NORMAL BATCH L5	52 psig	180 psig	232	110	2.4 mM	0.4mM	0.2M	0.72	0.98

13. L1 Visualization

In this section, we examine the heatmaps for L1 colored by the top 3 process parameters as indicated by Shapley analysis presented in the main text. As illustrated in **Fig. 8**, the high normal regioselective conditions are generally found at the low end of the reaction temperature range, while the iso-selective conditions are found at elevated temperatures.

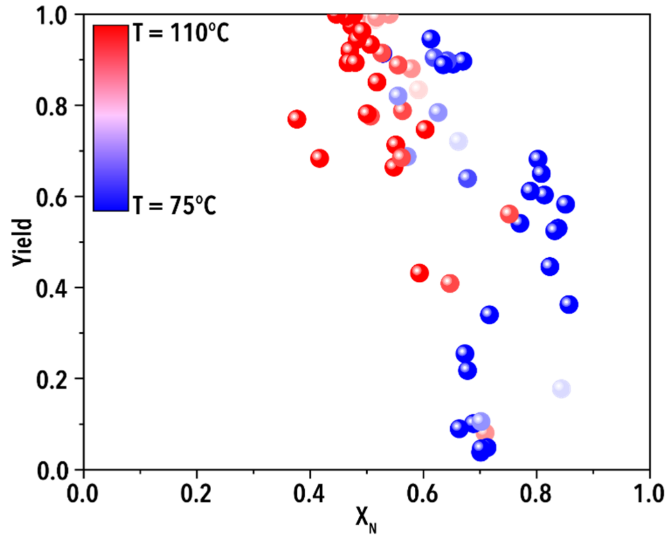

Figure 8. *L1 Pareto heatmap of reaction temperature.*

As presented in **Fig. 9**, the high normal regioselective conditions are generally found at high ligand to rhodium ratios while the iso-selective conditions are found at lower ligand to rhodium ratios.

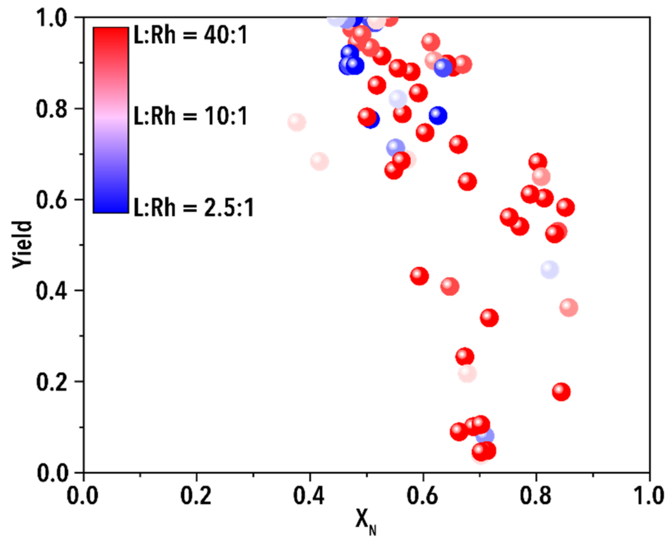

Figure 9. *L1 Pareto heatmap of ligand to rhodium ratio.*

As shown in **Fig. 10**, both the Normal and Iso selectivity campaigns resulted in primarily higher H_2 flowrates with the low H_2 flowrates occurring at low reaction yields, indicating the increased activity at elevated partial pressures of hydrogen.

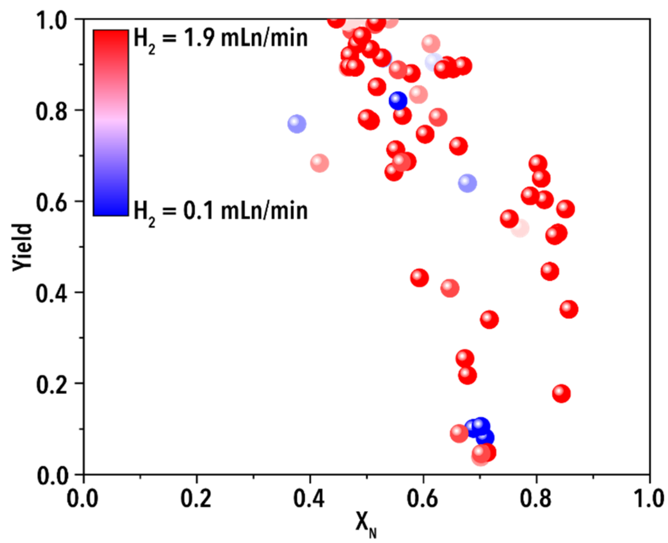

Figure 10. L1 Pareto heatmap of hydrogen flowrate.

14. Digital Twin Architecture

In this section, we discuss the steps to use the experimental data generated by Fast-Cat and build a rigorous machine learning (ML) model (*i.e.*, digital twin) based on deep neural networks (DNNs). The digital twin is built to effectively learn and analyze the complex reaction space of interest using autonomously generated experimental data by Fast-Cat. The coefficient of determination, also known as R^2, represents the proportion of the variance in the dependent variable that can be explained by the independent variables in a statistical model. Here, we use R^2 value of the ML model validation parity plot as a performance metric of each trained NN model architecture.

The programming and model training were carried out in Python 3.9.13, leveraging essential libraries including numpy, pandas, and tensorflow. In the first step, we compared the performance of a feedforward *vs*. a cascade forward NN (**Fig. 11**) for modelling the hydroformylation data generated by Fast-Cat. The goal of NN architecture search was to identify the optimal NN structure, including the most optimal number of layers and nodes, by maximizing the coefficient of determination. **Fig. 12A** and **12B** present the R^2 values of aldehyde yield and regioselectivity in the presence of ligand L1 obtained on the full dataset for a single feedforward and cascade feed forward NN model with *n* (x-axis) layers, each consisting of *m* (y-axis) nodes,

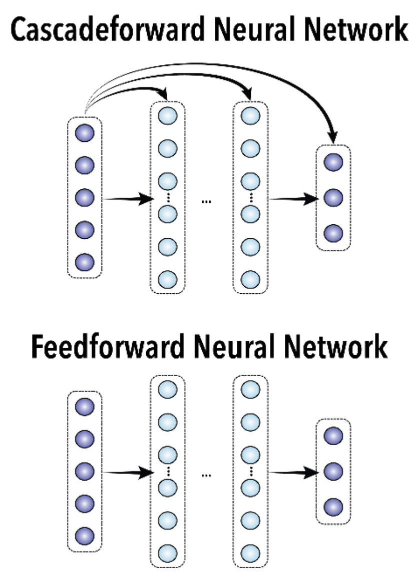

Figure 11. Comparison between feedforward and cascade forward structure.

respectively. Overall, as shown in **Fig. 12** the cascade NN architecture resulted in better performance compared to the feedforward NN architecture. The enhanced performance of the cascade NN architecture can be attributed to its enhanced multi-stage feature extraction, error propagation, handling of high-dimensional data, non-linear mapping capabilities, and robustness to noisy data. This is due to the multi-stage process, which allows the cascade NN to focus on different aspects of the data, enhancing its ability to represent complex relationships between input and output variables. The multi-stage architecture also allows for better error propagation and learning, as errors made in the initial stages can be further refined in subsequent stages. The cascading nature of the architecture also contributes to better robustness against noisy data, as it filters out noise and irrelevant features, focusing on the most salient information in the data.

In order to introduce uncertainty to the digital twin model of the hydroformylation reactions, we implemented an ensemble NN strategy. The suitable cascade NN structure range (layers and nodes) for each ligand/catalyst system tested by Fast-Cat was determined based on a structure screening study similar to the one performed for ligand L1 (**Fig. 12B**). The digital twin model was then constructed by randomly selecting the number of layers and nodes within the suitable range of layers and nodes obtained in the previous step. To improve the ML model performance, an outer NN with two layers was implemented to give a combination of each individual model's output instead of only averaging the outputs linearly[2]. Next, we studied the role of ensemble size on the digital twin performance. Since the ensemble NN model is constructed by randomly selecting the cascade NN structure within a specific range, conducting a direct one-to-one comparison may not provide a clear understanding of the effect of ensemble size on the accuracy of the digital twin. To address this issue, we used an average R^2 value of 50 generated ensemble NNs of the same size as the comparison metric to better understand the role of ensemble size on the digital twin performance. **Fig. 13** demonstrates that an ensemble size of 10 randomly selected cascade NNs achieves a satisfactory performance of R^2>0.7 using ligand L1 experimental data generated by Fast-Cat.

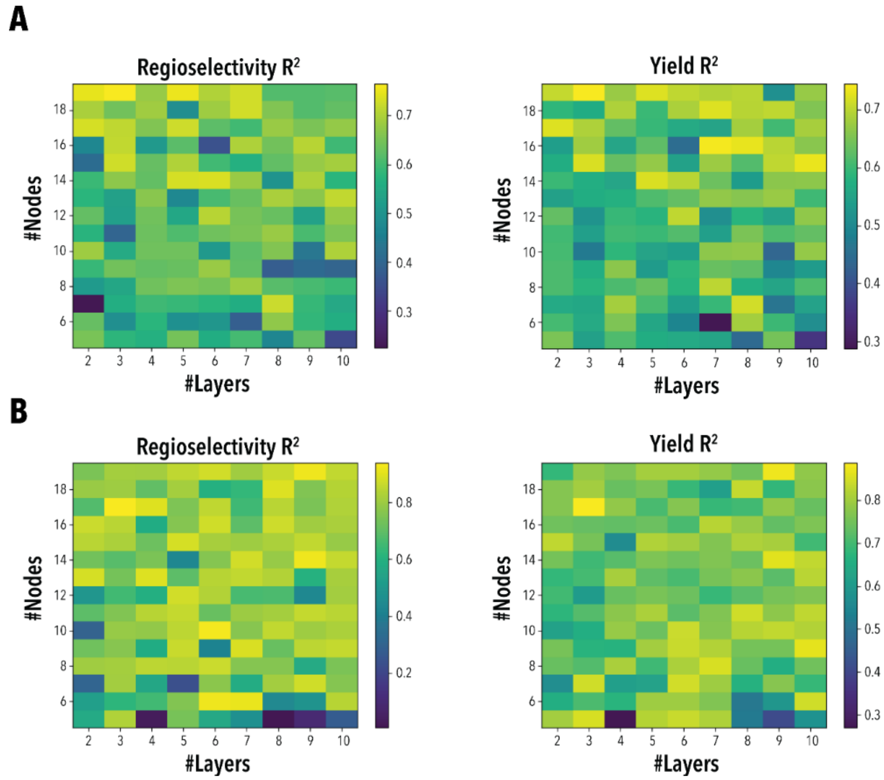

Figure 12. Digital twin model architecture study on the A) feedforward vs. B) cascade forward NN.

Increasing the ensemble size further would lead to longer training times without any gained benefit.

In order to further fine-tune the digital twin's performance, several other hyperparameters, including the learning rate and optimizer, were adjusted to achieve a smooth learning curve (see **Fig. 14**). Given that this is a regression problem, the loss function was defined as the root-mean-squared error, which the NN model seeks to minimize during training. *RMSprop* optimizer was used for optimization. The rectified linear unit (ReLU) activation function was applied to all hidden layers, while the output layers utilized a linear activation function. To ensure meaningful predictions, a threshold was introduced on the output layer, taking into account that the fractional aldehyde yield and regioselectivity values cannot exceed one. This measure was implemented to prevent the model from generating predictions that surpass the feasible experimental range and to maintain the physical constraints of Fast-Cat's hardware. Subsequently, the experimental data generated by Fast-Cat was utilized to build the digital twin model for each catalyst/ligand system using an 80% training data set, with 10% allocated for validation and another 10% for testing. The loss function of the validation set was monitored using an early-stopping function to prevent overfitting. The test data set, which the model has never encountered before, was used to assess the ML model's performance in the reaction space. To demonstrate the digital twin's performance, the average of model predicted values on the dataset *vs*. the actual values from the experiment are depicted **Fig. 15**.

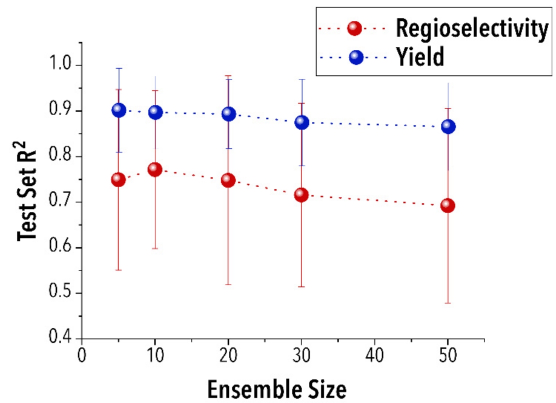

Figure 13. *Average performance of an ensemble of 50 randomly selected cascade NNs of a size 5, 10, 20, 30, and 50.*

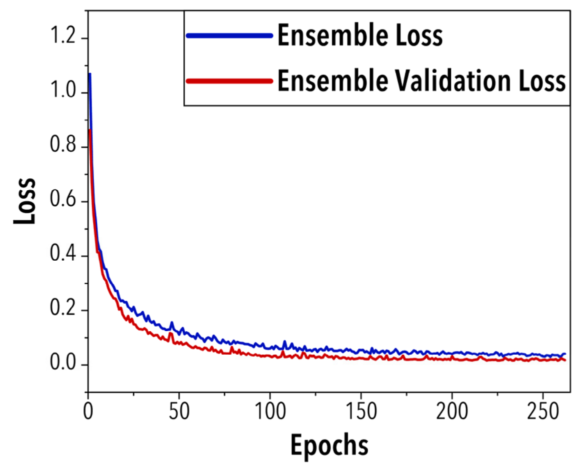

Figure 14. *The learning curve illustrates the digital twin model's training progress across epochs as it minimizes the mean squared error loss function. This curve demonstrates how the model's performance improves over time as it learns from the training data.*

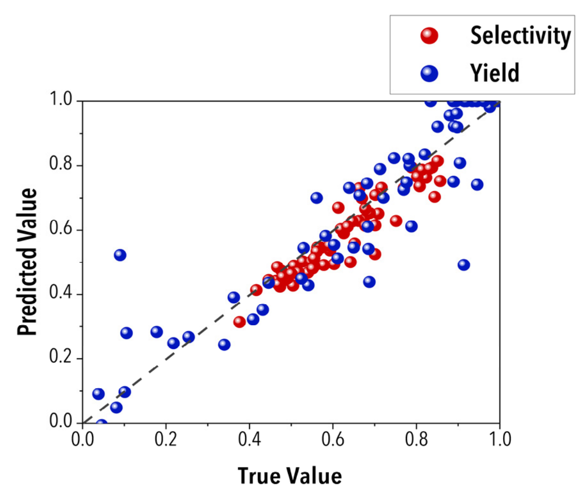

Figure 15. *Predicted values by the digital twin vs. the actual values from the experiment for ligand L1.*

Next, we performed the same hyper-parameter screening and optimization to build digital twins of other ligands (L2-L6). The parity plots and R^2 values for ligands L2 to L6 are provided (**Fig. 16, Table 12**). However, there is a notable distinction in the case of L4 and L5, as they exhibited a higher degree of linear regioselectivity. Consequently, the regioselectivity remained relatively unchanged over a wide range. This observation is illustrated in **Fig. 16**, where a cluster of data points is evident on the parity R^2 plot. For such clusters, R^2 may not serve as an adequate metric due to the fact that an infinite number of lines can pass through a cluster of points. Therefore, in this scenario, we utilized the root-mean-squared-error (RMSE) as a metric to fine-tune the digital twin for regioselectivity.

Table 12. *R^2 values on the full dataset for digital twins of ligands L1-L5.*

Ligand \ R^2	**Yield**	S_N
L1	0.93	0.94
L2	0.79	079
L3	0.95	0.74
L4	0.79	0.75
L5	0.94	0.87
L6	0.76	0.78

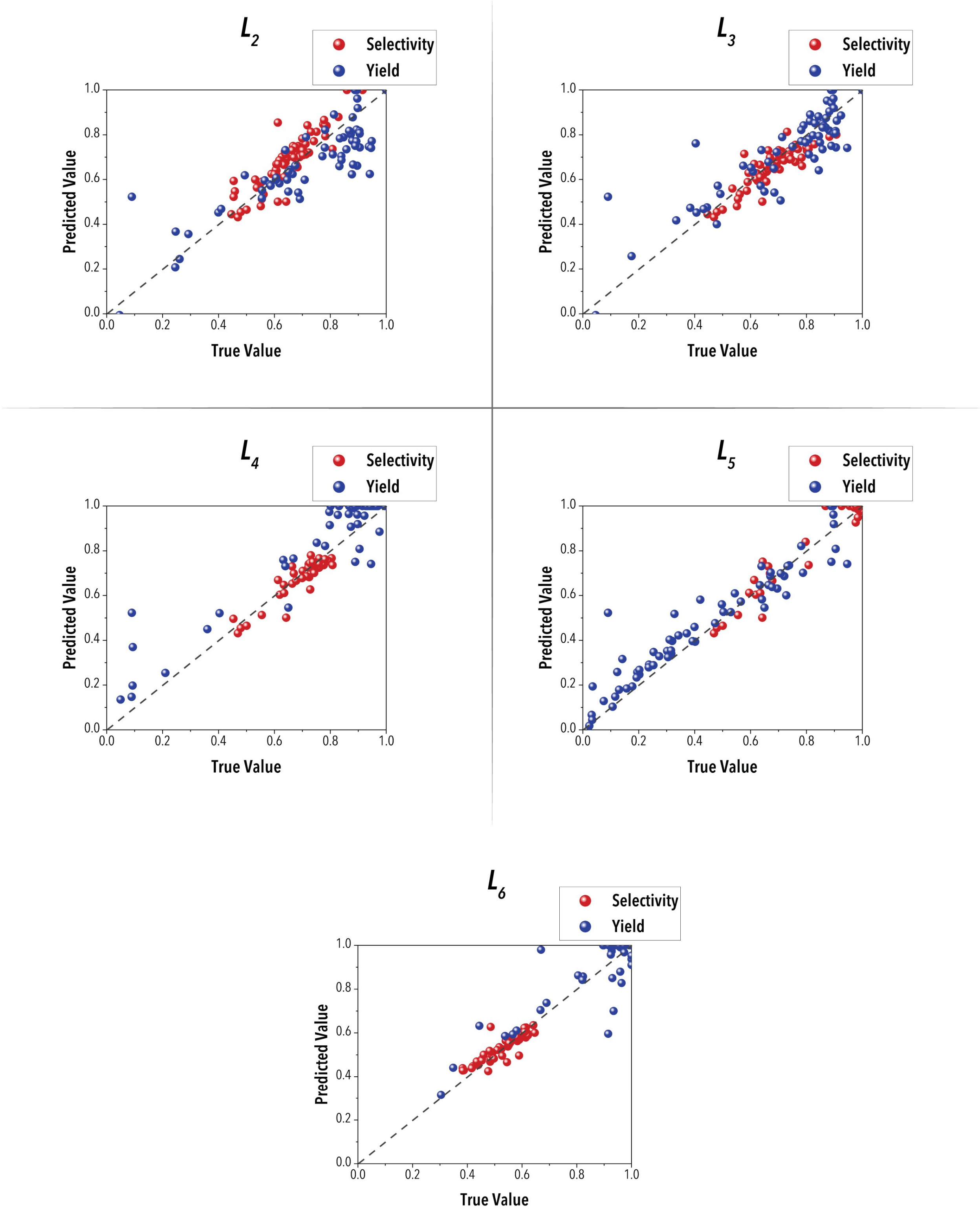

Figure 16. *Predicted values by the digital twin vs. the experimental values from the experiment for ligands L2-L6.*

15. Digital Twin Feature Analysis

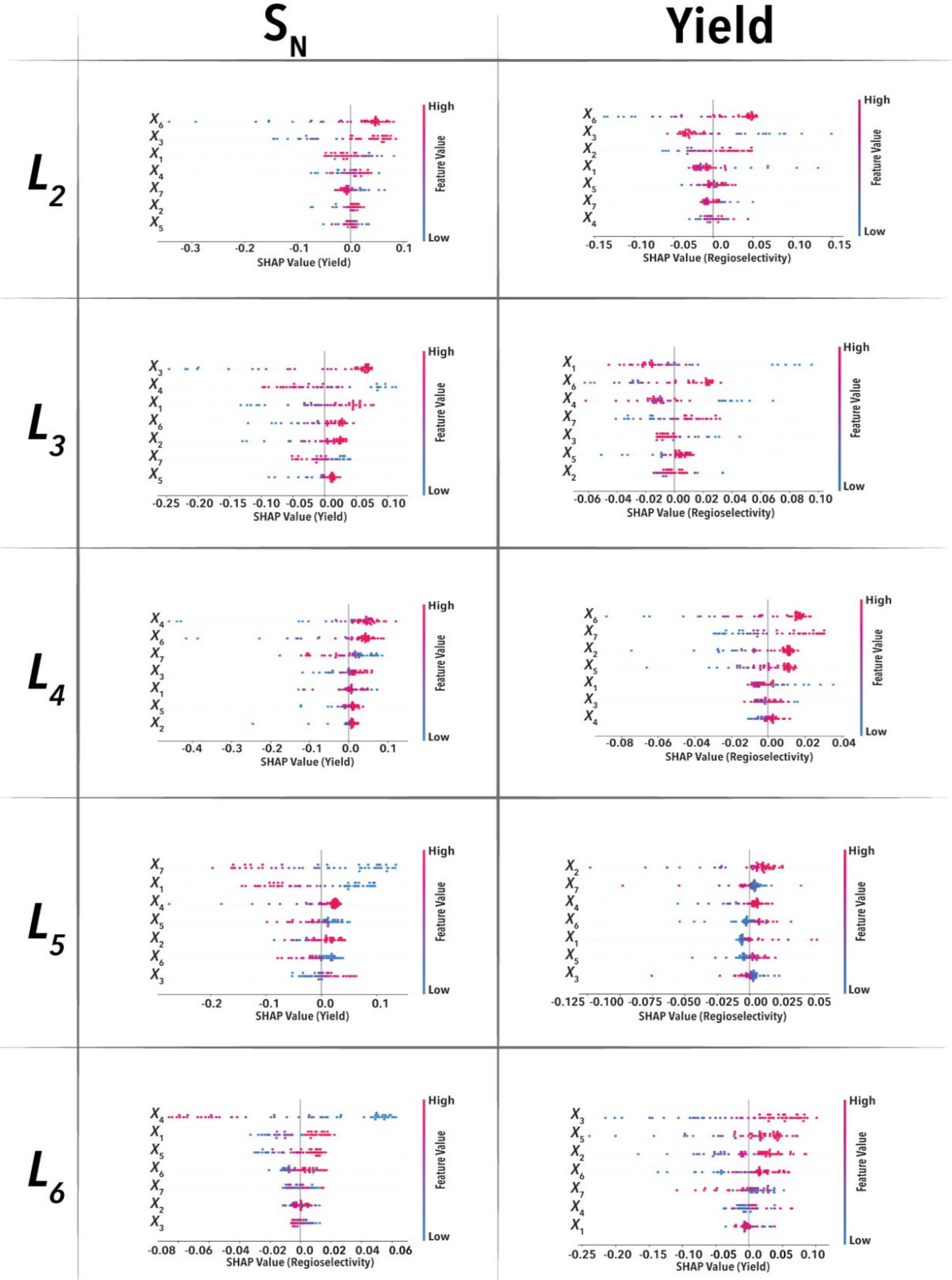

Figure 17. Shapley analysis of ligands L2-L6 using digital twins built on the experimental data generated by Fast-Cat. The Shapley analyses demonstrate differences in the significance of individual model features on aldehyde yield and regioselectivity across different ligands tested by Fast-Cat. The model feature variations can be linked to inherent structural distinctions that may affect the binding with Rh and consequently influence the underlying reaction mechanism.

Once we obtained a fine-tuned digital twin for each tested ligand by Fast-Cat, we conducted Shapely analysis on each digital twin model to assess how different model features influence the outcomes of the hydroformylation of 1-octene. The summarized results of these analyses are presented in (**main text Fig. 5C**).

Upon examining the Shapely plots for ligands L2-L6 (**Fig. 17**), it becomes evident that the general role of the hydroformylation process parameters is consistent across all tested ligands. For instance, increasing the reaction temperature generally leads to higher aldehyde yield but lower regioselectivity across all ligands. However, it is worth noting that the significance of each feature varies depending on the ligand's structure and steric hindrance.

Ligand L2 was selected as the most similar ligand to L1 based on its favorable attributes, including specific coordination properties, chelating ability, and significant steric and electronic effects. These properties play a crucial role in determining the stability and the catalyst regeneration. Ligand fraction and pressure are both the most critical features affecting total aldehyde yield and regioselectivity. Additionally, the flow rates of CO and H_2 are among the top three critical factors influencing the Fast-Cat's outputs.

Similarly, fluorophosphite ligand L3 demonstrates similar behavior to ligands L1 and L2. Notably, the flow rate of CO has a significant impact on both the aldehyde yield and regioselectivity. On the other hand, while the ligand fraction also plays a crucial role, its primary influence is observed in affecting the aldehyde regioselectivity.

The reason for ligand fraction being the most important feature in the hydroformylation of 1-octene, particularly in influencing aldehyde regioselectivity, can be attributed to the lower steric bulk of L2 and L3 compared to other ligands. This lower steric bulk allows L2 and L3 to form bis-complexes more favorably with the Rh catalyst. As a result, the relative ratios of ligand to Rh and CO flow rate play a more significant role in determining the overall direction of the aldehyde regioselectivity.

Triphenylphosphine (L4) is a low steric bulk ligand with smaller groups surrounding the central phosphorus atom. It can effectively coordinate with the Rh, forming stable complexes during the hydroformylation process. Therefore, the olefin fraction mainly affects the aldehyde yield. However, L4's regioselectivity is unique, as its ligand:Rh ratio significantly impacts the overall hydroformylation reaction. The ligand:Rh ratio influences the catalyst's geometry and electronic properties, affecting the reaction pathway and regioselectivity. L4's narrow range of aldehyde regioselectivities favors a specific regioisomer, resulting in a limited variety of aldehyde products.

Similar to L4, Xantphos (L5) is a mostly linear-favored ligand due to the low steric hinderance, and the regioselectivity will not change dramatically by changing the reaction features. H_2 flowrate and olefine fraction are mainly influencing the aldehyde yield. Since Xantphos is a bidentate phosphine ligand, the H_2 flowrate emerges as the most crucial factor affecting the overall hydroformylation of 1-octene in the presence of Xantphos. Xantphos exhibits excellent regioselectivity for the linear aldehyde product but shows relatively low activity in promoting the reaction. The bidentate nature of Xantphos allows it to form bis-complexes with the Rh catalyst, reducing the significance of the ligand:Rh ratio. Consequently, controlling the H_2 flowrate becomes vital as increased H_2 partial pressure promotes the reaction rate by providing essential hydrogen atoms for the catalytic cycle.

Next, the digital twin of each ligand built using the experimental data generated by Fast-Cat was utilized to visualize the hydroformylation of 1-octene in the presence of the corresponding ligand. To better understand the role of the top three model features (process parameters) from the Shapely analysis for ligands L2-L6, corresponding surface plots (reaction visualization) are presented in **Figs. 18-22**.

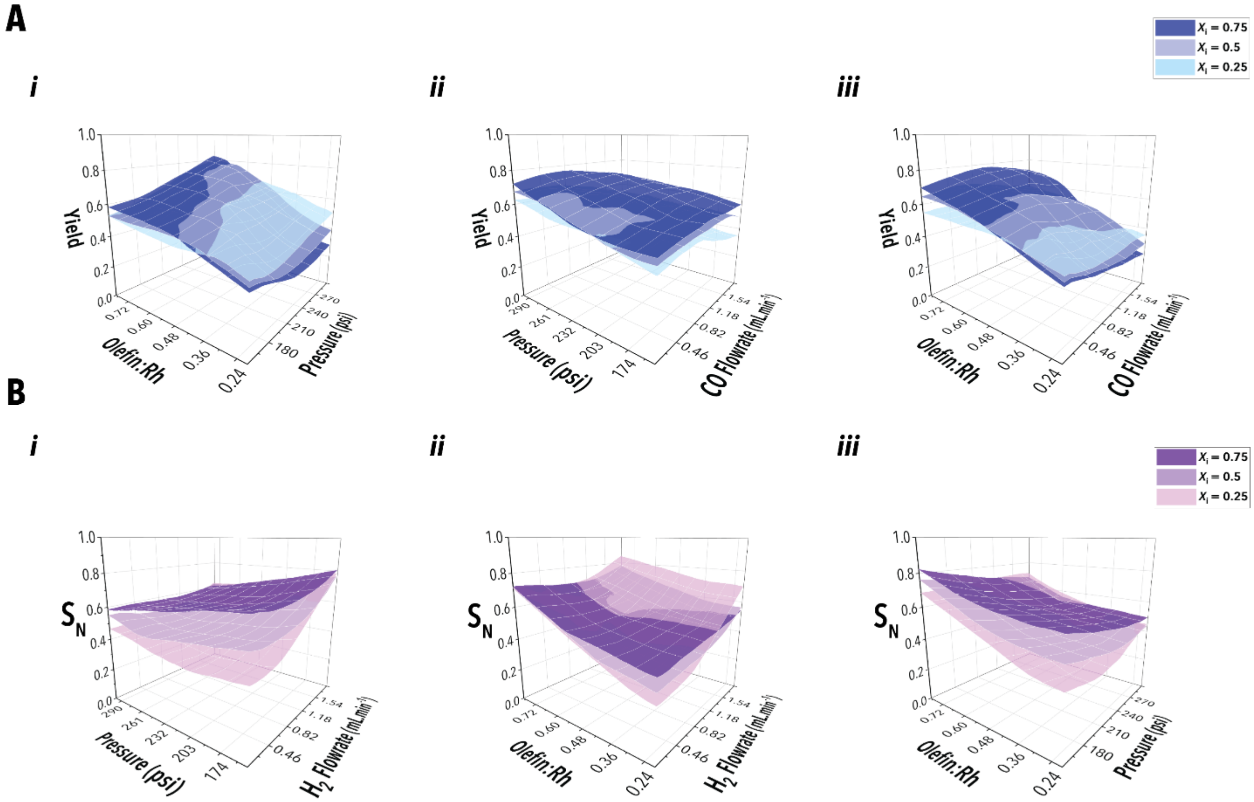

***Figure 18.** The hydroformylation reaction space exploration of 1-octene with ligand L2 using the digital twin built by the in-house experimental data of Fast-Cat. (A) Surface plots of predicted aldehyde yield with L2 as a function of (i) pressure and olefin to Rh fraction, (ii) pressure and CO flowrate, and (iii) CO flowrate and olefin to Rh fraction. (B) Surface plots of predicted aldehyde regioselectivity with L2 as a function of (i) Pressure and H_2 flowrate, (ii) olefin to Rh fraction and H_2 flowrate, and (iii) olefin to Rh fraction and pressure. For all surface plots shown in panels A and B, the process parameters other than the two variable parameters used for each surface plot were maintained at a constant value (X_i) for each surface.*

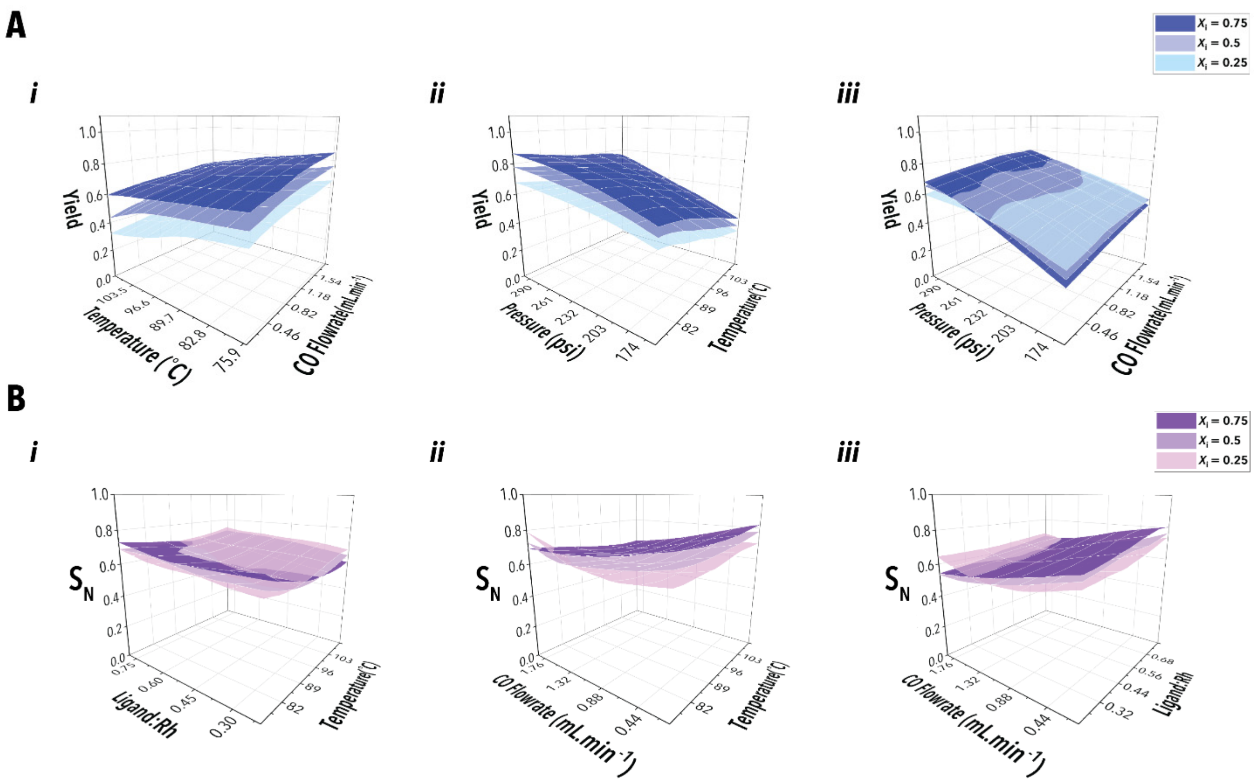

***Figure 19.** The hydroformylation reaction space exploration of 1-octene with ligand L3 using the digital twin built by the in-house experimental data of Fast-Cat. (A) Surface plots of predicted aldehyde yield with L2 as a function of (i) temperature and CO flowrate (ii) pressure and temperature, and (iii) CO flowrate and pressure. (B) Surface plots of predicted aldehyde regioselectivity with L3 as a function of (i) ligand to Rh fraction and temperature, (ii) CO flowrate and temperature, and (iii) ligand to Rh fraction and CO flowrate. For all surface plots shown in panels A and B, the process parameters other than the two variable parameters used for each surface plot were maintained at a constant value (X_i) for each surface.*

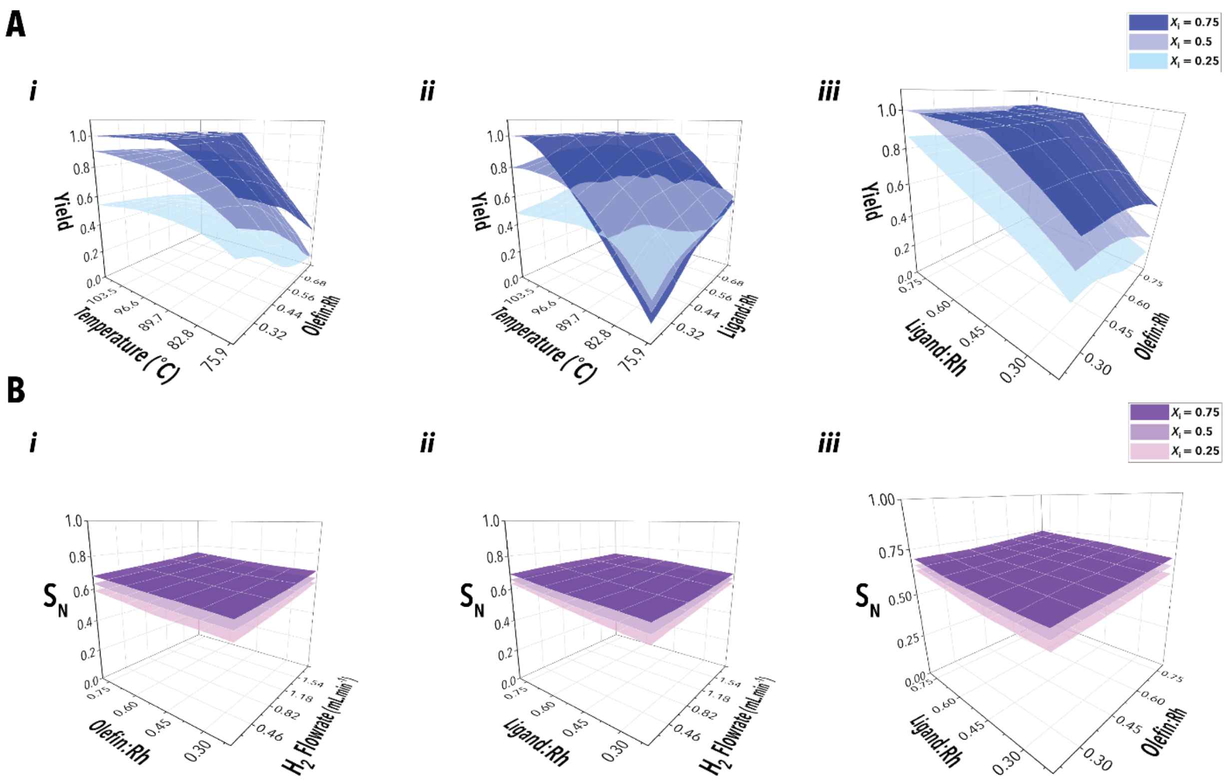

***Figure 20.** The hydroformylation reaction space exploration of 1-octene with ligand L4 using the digital twin built by the in-house experimental data of Fast-Cat. (A) Surface plots of predicted aldehyde yield with L4 as a function of (i) temperature and olefin to Rh fraction, (ii) temperature and ligand to Rh fraction, and (iii) ligand to Rh fraction and olefin to Rh fraction. (B) Surface plots of predicted aldehyde regioselectivity with L4 as a function of (i) olefin to Rh fraction and H_2 flowrate, (ii) ligand to Rh fraction and H_2 flowrate, and (iii) olefin to Rh fraction and ligand to Rh fraction. For all surface plots shown in panels A and B, the process parameters other than the two variable parameters used for each surface plot were maintained at a constant value (X_i) for each surface.*

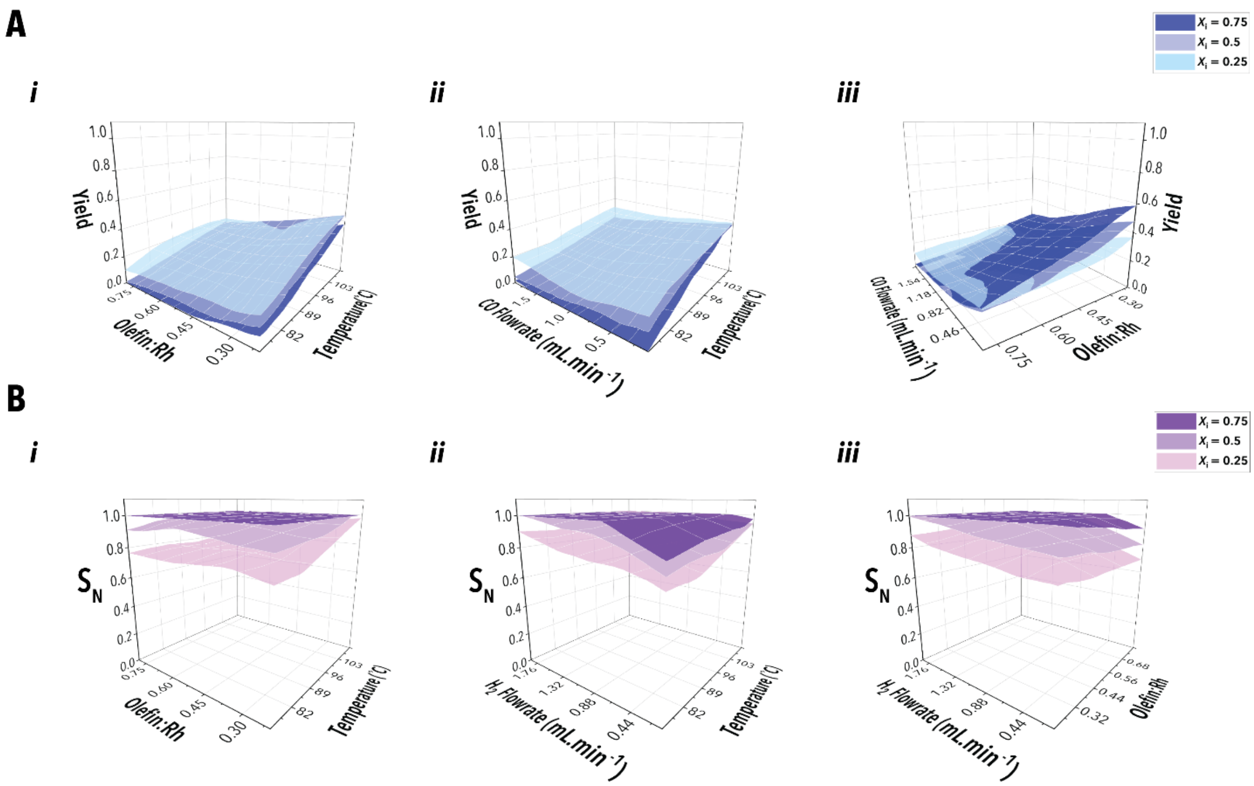

Figure 21. *The hydroformylation reaction space exploration of 1-octene with ligand L5 using the digital twin built by the in-house experimental data of Fast-Cat. (A) Surface plots of predicted aldehyde yield with L4 as a function of (i) olefin to Rh fraction and temperature, (ii) temperature and CO flowrate, and (iii) Co flowrate and olefin to Rh fraction. (B) Surface plots of predicted aldehyde regioselectivity with L5 as a function of (i) olefin to Rh fraction and temperature, (ii) temperature and H_2 flowrate, and (iii) olefin to Rh fraction and H_2 flowrate. For all surface plots shown in panels A and B, the process parameters other than the two variable parameters used for each surface plot were maintained at a constant value (X_i) for each surface.*

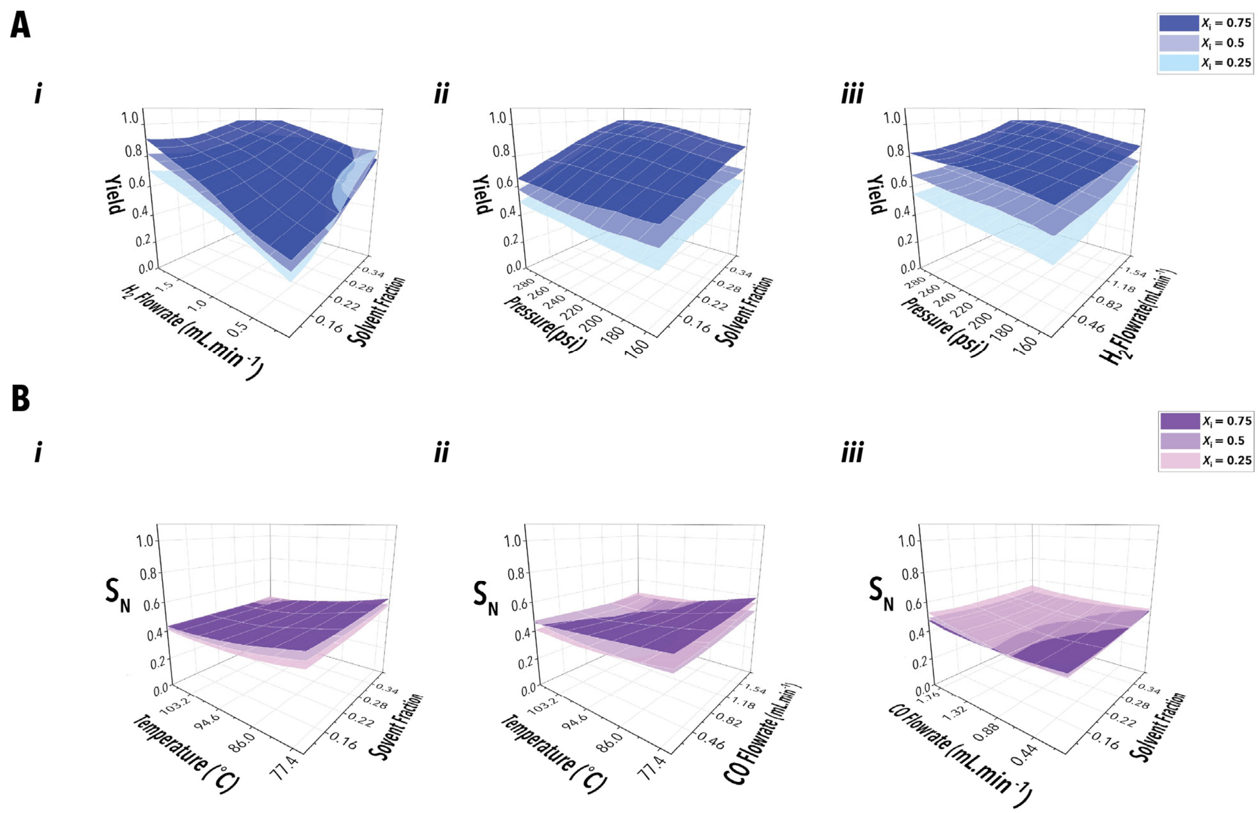

Figure 22. *The hydroformylation reaction space exploration of 1-octene with ligand L6 using the digital twin built by the in-house experimental data of Fast-Cat. (A) Surface plots of predicted aldehyde yield with L4 as a function of (i) H_2 flow rate and solvent fraction, (ii) pressure and solvent fraction, and (iii) pressure and H_2 flowrate. (B) Surface plots of predicted aldehyde regioselectivity with L6 as a function of (i) solvent fraction and temperature, (ii) temperature and CO flowrate, and (iii) solvent fraction and CO flowrate. For all surface plots shown in panels A and B, the process parameters other than the two variable parameters used for each surface plot were maintained at a constant value (X_i) for each surface.*

Next, in order to illustrate the uncertainty of the trained digital twin in predicting the hydroformylation reaction outcomes, the standard deviation of the ensemble NN predictions for each input reaction condition was utilized to project the upper confidence and lower confidence of the digital twin's predictions for L1 (**Fig. 23**).

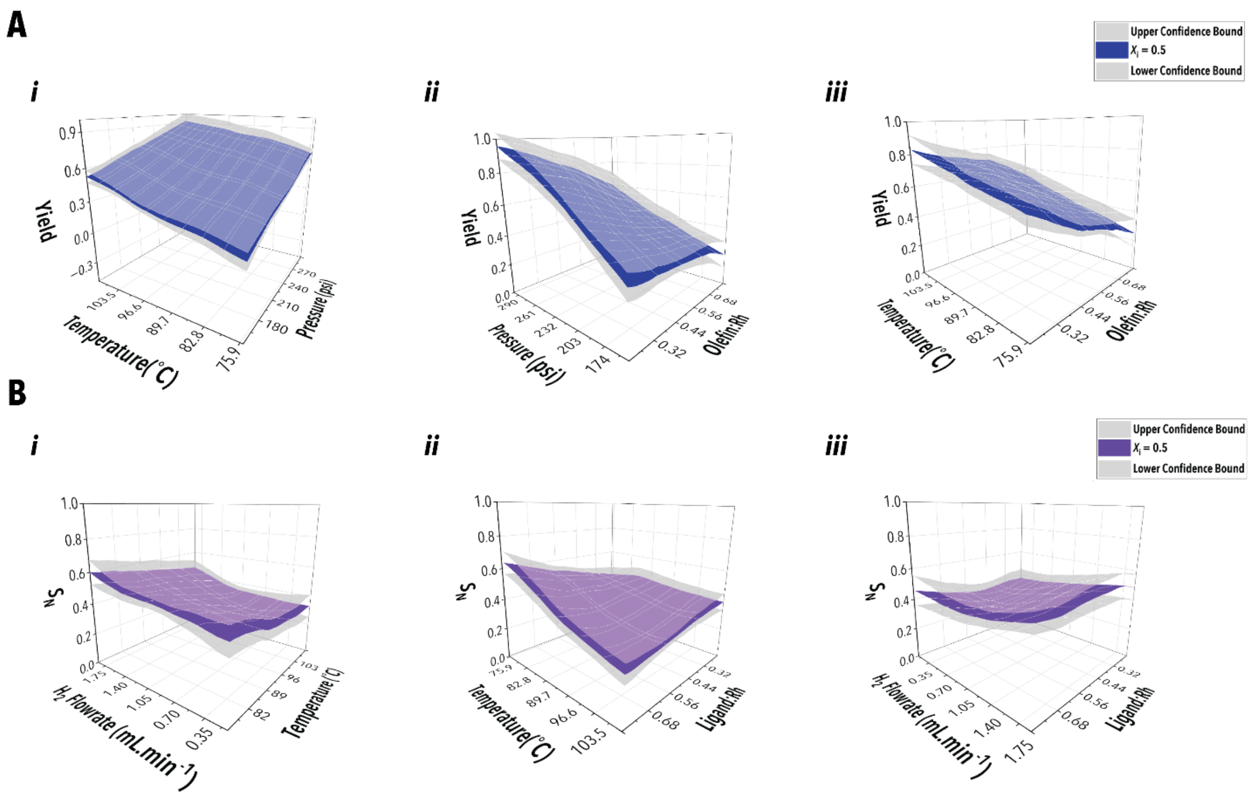

Figure 23. The hydroformylation reaction space exploration of 1-octene by considering the model uncertainty with ligand L1 using the digital twin built by the in-house experimental data of Fast-Cat. (A) Surface plots of predicted aldehyde yield with L1 as a function of (i) pressure and temperature, (ii) pressure and olefin to Rh fraction, and (iii) temperature and olefin to Rh fraction. (B) Surface plots of predicted aldehyde regioselectivity with L1 as a function of (i) temperature and H_2 flowrate, (ii) temperature and ligand to Rh fraction, and (iii) ligand to Rh fraction and H_2 flowrate. For all surface plots shown in panels A and B, the process parameters other than the two variable parameters used for each surface plot were maintained at a constant value (X_i) for each surface. Upper and lower confidence bound are respectively the final output of the digital twin plus and minus the standard deviation of the outputs of each individual model prediction in the ensemble.

S16. Supplementary References

1. Ibrahim, M. Y. S., Bennett, J. A., Mason, D., Rodgers, J. & Abolhasani, M. Flexible homogeneous hydroformylation: on-demand tuning of aldehyde branching with a cyclic fluorophosphite ligand. *Journal of Catalysis* **409**, 105-117 (2022). https://doi.org:https://doi.org/10.1016/j.jcat.2022.03.030
2. Shahhosseini, M., Hu, G. & Pham, H. Optimizing ensemble weights and hyperparameters of machine learning models for regression problems. *Machine Learning with Applications* **7**, 100251 (2022). https://doi.org:https://doi.org/10.1016/j.mlwa.2022.100251