A self-driving laboratory optimizes a scalable process for making functional coatings

Supplementary information

Connor C. Rupnow[1], Benjamin P. MacLeod[2], Mehrdad Mokhtari[2], Karry Ocean[2], Kevan E. Dettelbach[2], Daniel Lin[2], Fraser G. L. Parlane[2], Hsi N. Chiu[2], Michael. B. Rooney[2], Chris E. B. Waizenegger[2], Elija de Hoog[2], Curtis P. Berlinguette[1-4,*]

[1]Department of Chemical & Biological Engineering, The University of British Columbia, 2360 East Mall, Vancouver, British Columbia, Canada V6T 1Z3

[2]Department of Chemistry, The University of British Columbia, 2036 Main Mall, Vancouver, British Columbia, Canada V6T 1Z1

[3]Stewart Blusson Quantum Matter Institute, The University of British Columbia, 2355 East Mall, Vancouver, British Columbia, Canada V6T 1Z4

[4]Canadian Institute for Advanced Research (CIFAR), MaRS Innovation Centre, 661 University Ave. Suite 505, Toronto, Ontario, Canada M5G 1M1

*Email: cberling@chem.ubc.ca

Outline

Supplementary discussion

Sources of variability in the experiment

Across all experimental conditions tested during the optimization campaign shown in **Fig. 2** of the main text, the mean difference in conductivity of duplicates was 0.32 MS/m across all the samples. When experimental conditions yielding zero conductivity were excluded, the mean difference in conductivity of duplicates was 0.43 MS/m. There are several possible sources of conductivity variations between duplicate samples in this experiment. These sources include: variations in the precursor ink composition due to finite pipetting accuracy (see **Figure S11**); variable thermal contact of the glass substrates with the heated spraycoater fixture; and, variability in the time each sample spent on the heated fixture after the completion of spraycoating due to the variations in the schedule of the central robotic arm.

The model predictions improved over time

Cross sections of the 8-dimensional gaussian process model can be seen in **Figure S12**. During the optimization, the model r^2 value (a measure of fit quality) was saved every iteration. The model fit improved over the course of the campaign, converging to a final r^2 value of 0.95 (see **Figure S13**). The model was evaluated using leave-one-out cross-validation (see **Figure S14**) yielding an r2 value of 0.88.

Following the optimization, the conductivity values were compared with conductivity values predicted by the model. The GP model (see methods) was built over the data set iteratively, mimicking the order of the experiments. A four-sample delay was also added to mimic the parallelization (e.g. the model used to predict the conductivity of sample 45 only included information up to sample 41). We noted with interest that predictions significantly improved for experiments with beta values of 0.25 and 25 whereas the predictions got significantly worse for experiments chosen with beta value of 400 and space-filling mode (see **Figure S15**). This observation suggests that the model may have improved further had the optimization continued past 91 experiments.

Manipulated variables bounds

Four out of seven manipulated variables converged on values within 10% of the bounds (see **Fig. 2** of main text). This suggests that widening the bounds of the search space might yield films with even higher conductivities. Such a change, however, would have made the search space larger and thus potentially slower to search.

Discussion of the scale-up experiment

The conductivity of the large-scale sample is slightly higher than the small-scale sample likely for two reasons:

1. The larger coating has a more uniform coating due to the relative size difference between the spray shape and the substrate. The spray cone has an approximate diameter of ~16 mm which is greater than half the size of the 25 mm width substrate.
2. Four-point probe conductance measurements assume infinite size in the x-y direction. That means that conductance will measure higher for identical coatings with larger-areas.

The conductivity of the champion sample measured by Ada (see supplementary methods - autonomous workflow steps 4-6) is slightly different from the conductivity calculated using the manual data processing method (see supplementary methods - scale-up experiment). This is likely due to the autonomous measurements not being located at the most conductive point on the sample. The scale-up method, however, measures nearly the entire film and includes the most conductive locations in the measurements.

Supplementary tables

Table S1: Manipulated parameter bounds

input variable	definition	lower bound	upper bound	units
DMSO content	The amount of DMSO in the precursor ink as a volume fraction	0	0.3	v / v
precursor concentration	The amount of precursor (palladium nitrate and acetylacetone) in the ink as a mass concentration	10	20	mg / mL
ink flowrate	The rate at which the syringe pump ejects the ink out of the spray coater nozzle	2	8	µL / s
air flowrate	The amount the electronic proportional valve was open as a percentage of the maximum open state	65	100	control valve %
number of passes	The number of times the spray pattern is repeated for a given sample (see fig S10)	1	10	passes
height	The distance between the nozzle and the substrate	10	25	mm
hotplate temperature	The temperature of the surface of the aluminum hotplate fixture	220	300	°C

Supplementary materials and methods

Manual preparation of stock solutions

Stock solutions were prepared by hand and stored in capped 2-mL HPLC vials. These vials were placed in a tray that was accessible to a robotic arm of the self-driving laboratory. The $Pd(NO_3)_2•H_2O$ and acetylacetone solutions were prepared in acetonitrile (MeCN) at a concentration of 30 mg mL^{-1}. Palladium(II) nitrate hydrate ($Pd(NO_3)_2•H_2O$; Pd ~40% m/m; 99.9% Pd purity, CAS 10102-05-3) was purchased from Strem Chemicals, Inc. MeCN (CAS 75-05-8; high-performance liquid chromatography (HPLC) grade, ≥99.9% purity) and acetylacetone (CAS 123-54-6; ≥99% purity) were purchased from Sigma-Aldrich. All chemicals were used as received.

Manual preparation of consumables

The self-driving laboratory used the following consumables as received: glass substrates (75 mm × 25 mm × 1 mm microscope slides; VWR catalog no. 16004-430); 2-mL HPLC vials (Canadian Life Science); and 200-µL pipettes (Biotix, M-0200-BC). These items were placed in racks and trays for access by the self-driving laboratory.

Autonomous workflow step 1: select conditions for initial experiments

A set of 15 initial random experiments were chosen to initialize the optimization. This amount was chosen based on an arbitrary 2n+1 rule, where n in the number of dimensions. A random value between 0 and 1 was chosen for each experimental condition from a uniform random distribution (python numpy.random.random). This value was then scaled to the range of the variable. Since each experiment was performed in duplicate, 30 samples with 15 unique experimental conditions were used to initialize the optimization.

Autonomous workflow step 2: mix precursor

A 4-axis laboratory robot (N9, North Robotics) located within the precursor mixing station formulated each precursor ink by pipetting the stock solutions described above into a clean 2 mL HPLC vial. The total requested volume of the precursor ink 0.38 mL remained fixed while the volume of each individual stock solution within the precursor ink was allowed to vary. Each precursor ink had a requested precursor concentration (palladium nitrate and acetylacetone) and DMSO content. The required amounts of palladium nitrate and acetylacetone to reach the set concentration in 0.38 mL were first pipette into the vial. MeCN and DMSO were then added to the vial up to 0.38 mL in order to reach the set concentration and DMSO content. Gravimetric feedback from an analytical balance (ZSA120, Scientech) was used to record the true (or "realized") amount of dispensed precursor and minimize pipetting errors (see Fig. S11). The vial was then placed back into a holding tray once mixing was complete.

Autonomous workflow step 3: spray-coat

A blank glass substrate (75 mm × 25 mm × 1 mm microscope slides; VWR catalog no. 16004-430) and a vial containing the precursor ink were passed from the precursor mixing station to the spray-coating station. Glass substrates were stored on a substrate rack at the precursor mixing station. The N9 robot used a custom vacuum chuck to pick up a substrate from the substrate rack and drop it off at a transfer tray. This transfer tray allowed the transfer of samples between the N9 robot and a 6-axis robotic arm (UR5e, Universal Robotics). The UR5e would pick up the substrate from the transfer tray and drop it off at the spray-coating station. The UR5e uses a vacuum chuck affixed to the end effector, similar to the vacuum chuck that is used by the N9. A vial containing the precursor ink was picked up by the N9 robot and passed to the spray-coating station via the UR arm. The vials were transferred using a custom 3D-printed vial carrier that allowed the vials to be transported with the vacuum chuck on the UR robot arm.

The spray-coater was built from an ultrasonic nozzle (Microspray, USA) mounted to a custom motorized XYZ gantry system (Zaber Technologies Inc., Canada) above a hotplate (PC-420D, Corning, USA). A custom aluminum fixture (alloy 6061, 12.7 mm thick) with notches for substrate access was mounted to the hotplate. This fixture enabled substrates to be picked up and set down on the hotplate by the UR robot arm.

Precursor ink was front-loaded into the spray-coater nozzle via a needle that protrudes through the center bore in the ultrasonic nozzle. A pneumatic solenoid used compressed air to control the extension and retraction of the needle. A syringe pump (cavro centris pump PN: 30098790-B, Tecan Trading AG, Switzerland) controls the flow of precursor ink using isopropyl alcohol as a backing solvent. A 50 uL gap between the backing solvent and the precursor ink separated the two fluids in the tubing (Cole-Parmer, PTFE, 1 mm ID). The ultrasonic spray nozzle was operated at 3 W and 120 kHz. For each recipe, a total of 300 µL of precursor was sprayed onto a pre-heated glass substrate.

Hotplate temperature was measured using a thermocouple mounted to the surface of the hotplate aluminum fixture. The thermocouple would provide feedback to a PID controller (k_p=0.7, k_i=3.5, k_d=65) that would turn the hotplate on or off to maintain a constant surface temperature.

When spraying, the nozzle moved in a serpentine pattern consisting of twelve lines 50 mm in length and evenly spaced within the 25 mm wide glass slide (see Fig. S10a). This pattern was repeated for the requested number of passes with a 30 second delay between each pass. The spray-coater nozzle speed varied between 2 and 168 mm s^{-1}, depending on the number of passes. Nozzle speed was calculated by dividing the total distance that the nozzle would travel by the amount of time the nozzle would need to spray all 300 µL at a particular flowrate.

Compressed air was fed to the nozzle at 60 PSI but restricted by an electronic air control valve (EV-P-20-2550, Clippard, USA). This valve varied between 65 and 100 % of the maximum. A valve setting of 100% was measured to be 22 standard cubic feet per hour (SCFH) while a value of 65% was measured to be 16 SCFH. The height of the nozzle above the substrate, controlled by the motorized XYZ gantry system, varied between 10 and 25 mm. The spray flow rate, controlled by the syringe pump, varied between 2 and 8 µL s^{-1}. After spray-coating, the samples were left to rest on the hotplate for a minimum of 200 seconds.

Autonomous workflow step 4: XRF

Hyperspectral X-ray fluorescence (XRF) images of each sample were acquired using a Bruker M4 TORNADO X-ray fluorescence microscope. Samples were transported from the spray-coating station to the XRF microscope by the UR5e robot arm. The instrument was equipped with a customized sample fixture employing an alignment tool.

The XRF microscope operated using a rhodium X-ray source at 50 kV/600 µA/30 W. The X-ray optics yielded a 25 µm spot size on the sample. The instrument achieved an energy resolution of 10 eV via twin 30 mm^2 silicon drift detectors. Hyperspectral images were taken over a 16 mm × 16 mm area at a resolution of 40 × 40 pixels. The XRF spectra obtained (reported in counts) were scaled by the integration time (200 ms) and the energy resolution (10 eV) to yield units of counts s^{-1} eV^{-1}.

To quantify the relative amount of palladium in the film, the XRF spectra was integrated at the palladium Lyman-alpha X-ray fluorescence line (2.837 keV) from 2.6 to 3.2 keV. The resulting counts were converted to film thickness estimates by applying a calibration factor obtained using reference samples (see below). Five points of interest were defined within the XRF hypermap of the sample, spaced 2 mm apart vertically down the center of the XRF hypermap (see supplementary Fig. S10a). For each point of interest, the XRF counts per second were averaged over a 3 mm × 3 mm area.

Calibration of XRF signal against reference samples

To enable palladium film thickness to be estimated from the XRF signal, a calibration procedure was performed according to Macleod et al[15]. Sputtered palladium reference samples were fabricated having four different nominal thicknesses (10,

50, 100, and 250 nm). These samples were characterized by profilometry and XRF. A linear relationship between the film thickness and the XRF counts was observed with a slope of 3.658×10^{-4} nm cps^{-1} and intercept of -27.25 nm. This relationship was used to estimate the thickness of each sample from the XRF data.

Autonomous workflow step 5: Conductivity & Imaging

After XRF imaging, the UR5e robot transported the Pd film sample to the microscope and conductivity station. This station consists of a four-point conductivity probe and optical microscope that is serviced by a 6-degrees-of-freedom gantry robot comprised of motor stages from Zaber Technologies. A custom end-effector mounted on the motor stack is used to secure samples in place during characterization using spring-loaded tabs.

The samples were imaged using an Axio Imager Vario.Z2 microscope from Zeiss. Brightfield and darkfield images were captured at 2.5× and 20× magnification at 2 mm above and below the center of the film. Focus and exposure parameters for each sample were determined with an automatic routine through Zen Blue software provided by Zeiss.

Following optical microscopy, the gantry would move the sample to the four-point probe. Four-point probe conductance measurements were acquired using a Keithley Series K2636B System Source Meter instrument connected to a Signatone four-point probe head (part number SP4-40045TBN; 0.040-inch tip spacing, 45 g pressure, and tungsten carbide tips with 0.010-inch radii) by a Signatone triax to BNC feedthrough panel (part number TXBA-M160-M). The source current was stepped from 0 to 1 mA in 0.2 mA steps. After each current step, the source meter was stabilized for 0.1 s and the voltage across the inner probes was then averaged for three cycles of the 60 Hz power line (i.e., for 0.05 s) and recorded. Conductance measurements were made on the same five points of interest as analyzed in the XRF data (see Fig. S10a).

The samples were then positioned 30 cm below an optical camera (FLIR Blackfly S USB3; BFS-U3-120S4C-CS) using a Sony 12.00 MP CMOS sensor (IMX226) and an Edmund Optics 25 mm/f1.4 C-Series Fixed Focal Length Imaging Lens (#59–871) for imaging. White printer paper was placed 2.5 cm behind the sample to improve image contrast.

Autonomous workflow step 6: Process the raw data

The film conductivity was calculated using a custom data analysis pipeline implemented in Python using the open-source Luigi framework[42]. This pipeline combined conductance data and XRF data to estimate the film conductivity at each of the five points of interest on the sample.

For each set of current–voltage measurements at each position on each sample, a linear fitting algorithm was used to extract the conductance (dI/dV). The voltage compliance limit of the K2636B was set to 10 V and voltage measurements greater than 10 V were therefore considered to have saturated the Source Meter instrument and automatically discarded by the data analysis pipeline.

The conductivity of the thin films was then calculated by combining the 4-point-probe conductance data with the film thicknesses estimated by XRF:

$$\sigma = \frac{ln2}{\pi}\left(\frac{dI}{dV}\right)t^{-1} \quad (1)$$

where dI/dV is the conductance from the 4-point-probe measurement, t is estimated film thickness from the XRF measurements, and σ is conductivity.

Outliers were excluded from the conductance data using a kernel density exclusion method (see below). Outliers were also excluded from the XRF film thickness estimates using the same exclusion method. Conductivities were calculated for each position on the sample for which neither conductance nor XRF data was excluded. The mean of these conductivities was returned to the optimizer (see below). In cases where all points were discarded, a mean conductivity of 0 was reported.

The outlier kernel density exclusion method was performed by calculating Gaussian kernel density estimates for the conductance and XRF data, normalizing the density between 0 and 1, and rejecting data points with a kernel density below 0.3. Bandwidths of 5×10^{-3} $\mu\Omega^{-1}$ m^{-1} and 5×10^{3} cps were used for the conductance and XRF data, respectively.

Autonomous workflow step 7: Choose the next experiment

To choose the next experiment, Ada would request a new set of experimental conditions from a queue. If there was a sample in the queue, then Ada would make and test that sample. If the queue was empty, a new set of experimental conditions would be requested from the optimizer. Following the 15 random experiments (described in step 1), the optimizer was configured to pass experiments to the queue that were selected using Bayesian optimization.

The Bayesian optimization was performed using the Botorch python package[27] and would happen in two steps. First, a surrogate gaussian process (GP) regression model was built over all existing data. The data covariates were normalized to the unit cube and outcomes were standardized (zero mean, unit variance). A fixed noise GP model was used for the surrogate model with standard deviation of the input data estimated to be of 0.2 MS/m. We chose 0.2 MS/m as the noise estimate because it is 10% of the previously reported maximum conductivity value for this technique. All other hyperparameters for the GP model were default. The model predictions are discussed in the supplementary discussion and in Figs. S12-15.

Second, an acquisition function was selected and a new set of experimental conditions were acquired. The acquisition function would cycle between four modes; upper confidence bound (UCB) with three different beta values (beta of 0.25, 25, 400) and a space-filling point. The space-filling point would select an experimental condition in the parameter space that would maximize the distance to the closest other existing experiment. This cycling was necessary to accommodate our parallel workflow. The acquisition functions were configured to maximize the conductivity while also effectively searching the parameter space. Once a set of experimental conditions were found (via UCB or space-filling), the optimizer would duplicate the result and pass two sets of experimental conditions to the queue.

This autonomous workflow from steps 2 to 7 was repeated until a critical error ultimately stopped the optimization after 91 experiments. During the optimization, the self-driving laboratory was stopped briefly four times to refill consumables or to adjust a hardware component. Occasionally, these stops would result in some experimental conditions not having a duplicate (eg. experiments 44, 46, 91). Re-booting the self-driving laboratory after a brief stop would occasionally resulted in the same acquisition function to be used twice (or more) in a row (eg. experiments 16/17 and 60/61).

Space-filling point

The space-filling point algorithm employed a monte-carlo-based distance optimization. A selection of 100 random points were chosen in the parameter space. Each point was individually optimized according to a scheme which maximized the objective function, $\min(|| x - a ||^2)$, where x is the random point and a is a set of all available data points, subject to the parameter bounds (see Table S1). We used the scipy.optimize.minimize function to accomplish this maximization for each random point. This moves each point to the closest local maximum. The point that had the greatest objective function value (i.e. the point that was furthest away from the next closest point) was chosen as the next experiment to perform.

Scale-up experiment

Palladium precursor ink was sprayed on a 100 mm × 100 mm × 1.6 mm piece of glass. A new larger aluminum fixture (of the same thickness) was made for the hotplate to accommodate the larger piece of glass.The scale-up experiment used the same hotplate and thermocouple set-up as was used during the optimization. The volume of ink sprayed and the spray pattern were scaled-up while the experimental conditions remained the same as the optimization champion. The amount of ink sprayed increased from 0.3 to 2.4 mL and the pattern size increased from 25 mm × 50 mm with 12 evenly-spaced lines to 100 mm × 100 mm with 48 evenly-spaced lines (see pattern in SI **Fig. S10b**). The nozzle speed remained at 33.86 mm s^{-1} for both small and large spray patterns.

A small-scale sample from the champion experiment (sample 123 from experiment 61) was chosen to analyze for comparison against the large-scale sample. Conductance measurements were acquired with the four-point probe in a grid pattern on both samples. The small sample had a grid of 100 points extending 5 points spaced 4 mm in the x-direction and 20 points spaced 2mm apart in the y-direction. The large sample had a grid of 576 points extending 24 points spaced 4 mm apart in the x-direction and 24 points spaced 4mm apart in the y-direction. The samples were then both characterized for thickness. For the small sample, a high-resolution grid of XRF measurements was taken with each pixel spaced 0.2 mm apart and an integration time of 200 ms per pixel. For the large sample, a high-resolution grid of XRF measurements was taken with each pixel spaced 0.18 mm apart and an integration time of 200 ms per pixel. The high-resolution grid was down-sampled to match the pixel size and location of the conductance grid by averaging the high-res pixels in the bounding box of the conductance pixel size. The down-sampled XRF pixel map was converted to thickness using the same calibration curve as used during the optimization (see autonomous workflow step 4, calibration of XRF signal). The conductivity was calculated for each pixel using the same method as used during the optimization (see autonomous workflow step 6).

Supplementary figures

All figures containing numerical data were created in Python using the matplotlib library, with the exception of **Figure S8c** which was created using Microsoft Excel.

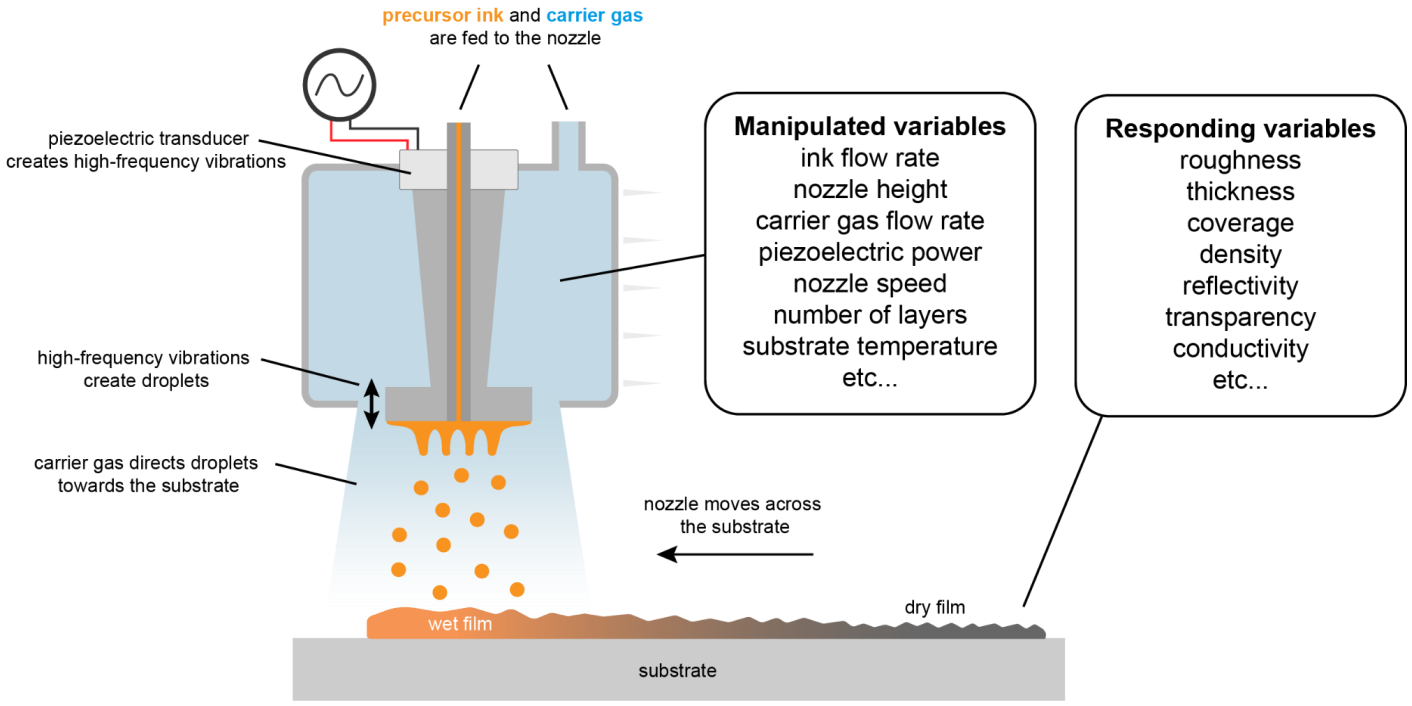

Figure S1 | Spray-coating is a scalable coating technique with numerous manipulated and responding variables. An ultrasonic nozzle moves in a pattern across a substrate. Precursor ink is fed via a syringe pump to the nozzle tip where an ultrasonic transducer breaks the fluid into tiny droplets. These droplets are then directed towards the substrate by a carrier gas. The droplets impact the substrate and coalesce to form a film. Spray-coating has many manipulated variables and is difficult to optimize. The manipulated variables must be tuned properly otherwise the resulting film will have poor responding properties.

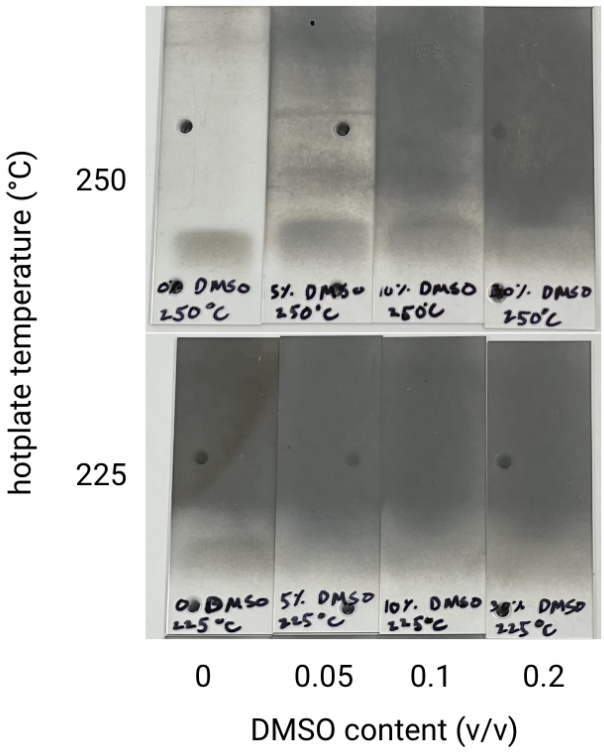

Figure S2 | Adding DMSO to an acetonitrile-based precursor ink enabled spray-coating at higher temperature. Two sets of spray-coated Pd films were prepared; the first at 225 °C and the second at 250 °C. Precursor inks containing MeCN and no DMSO (labelled "0% DMSO" on the left side) were unable to make continuous films at 250 °C. As higher relative amounts of DMSO was added to the precursor ink, continuous films were capable of being produced at 250 °C.

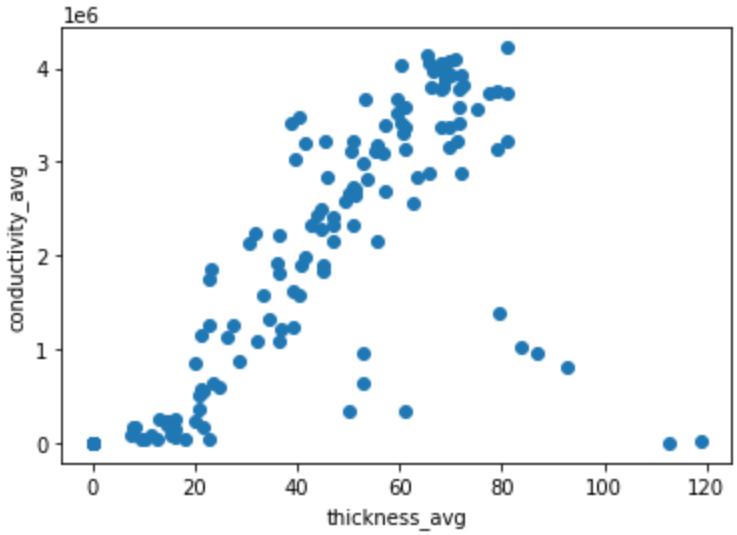

Figure S3 | Film conductivity is correlated with thickness for the spray-combustion synthesized films studied here. This plot shows the conductivity plotted against the thickness for all samples. There is a clear trend of increasing conductivity for increasing thickness, regardless of experimental conditions. The samples in the bottom right of the plot show that it is possible to achieve thick films with low conductivity. This is likely due to the temperature of the hotplate being insufficient to fully combust the precursors.

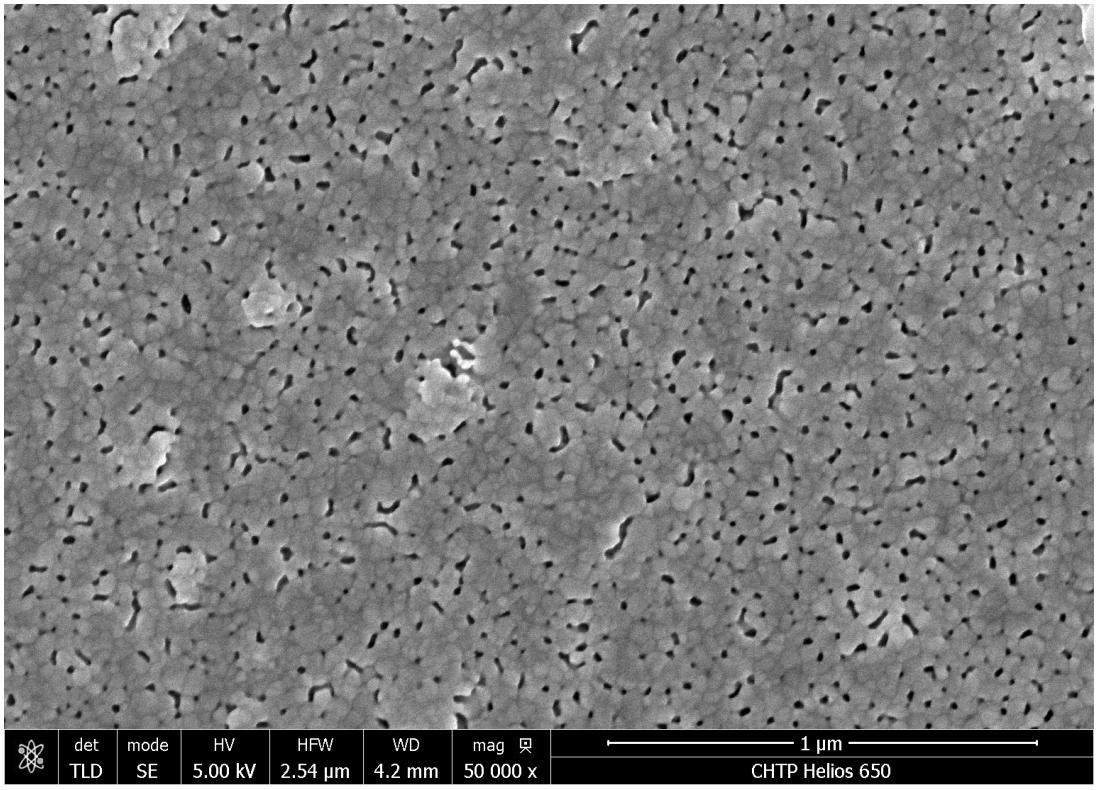

Figure S4 | Scanning electron microscopy of a sample made using the champion experimental conditions revealed a granular structure with pinholes. The grains are approximately 50 - 100 nm in diameter.

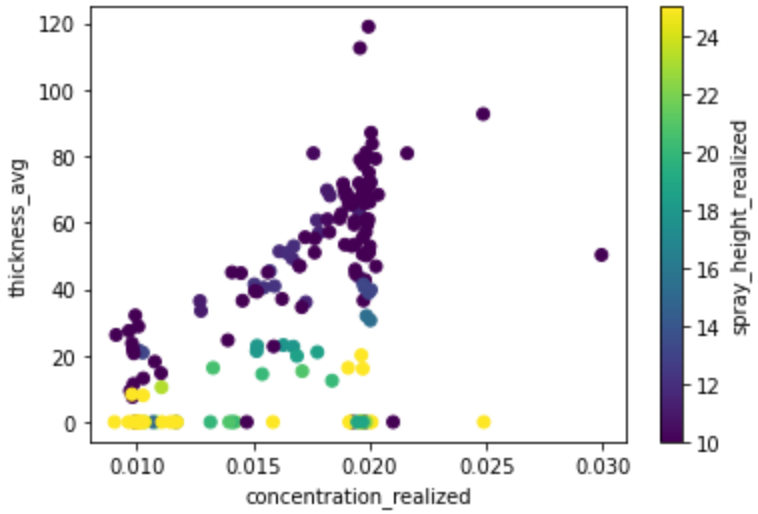

Figure S5 | Low spray height and high concentration were found to maximize Pd film thickness. This plot shows the thickness plotted against the realized concentration for all samples. As the concentration goes up, so does the maximum achievable thickness. The color indicates the height of the spraycoater nozzle. It can be seen that for the majority of samples low nozzle height produced films on the upper boundary of thickness, regardless of concentration. This indicates that low nozzle height improves the deposition efficiency.

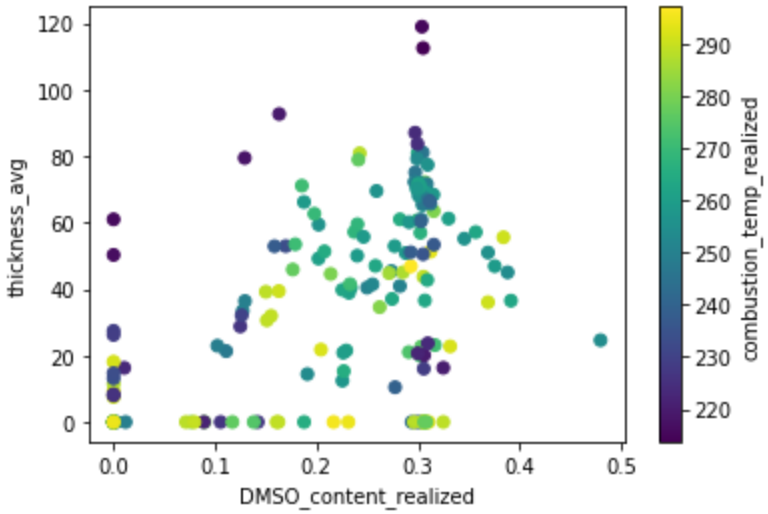

Figure S6 | Higher film thicknesses were associated with elevated DMSO content. This plot shows the film thickness against the realized DMSO content. It can be seen that thicker films are possible with higher DMSO content. The thickest films across DMSO content always have the lowest temperature (purple points) while the high temperature films (yellow points) are never on the maximum thickness boundary. It should be noted that thinner films across all DMSO amounts were the result of the effects of other experimental conditions (e.g. nozzle height, air flow rate, etc.).

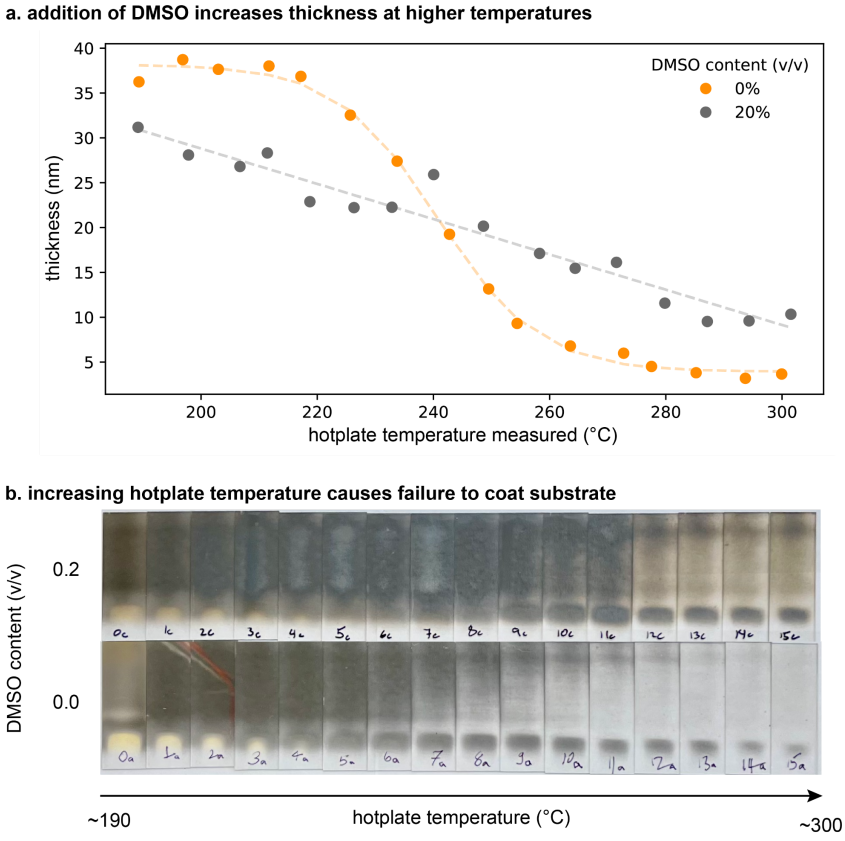

Figure S7 | Precursor inks containing DMSO yielded thicker films at higher temperatures. (a) Thickness measurements for films with no DMSO and with 20% DMSO across a range of temperatures. Lines were fit to the data to guide the eye. **(b)** The films from the above plot went through visible thickness changes at around ~235 °C for the samples without DMSO and ~270°C for the films with 20% DMSO.

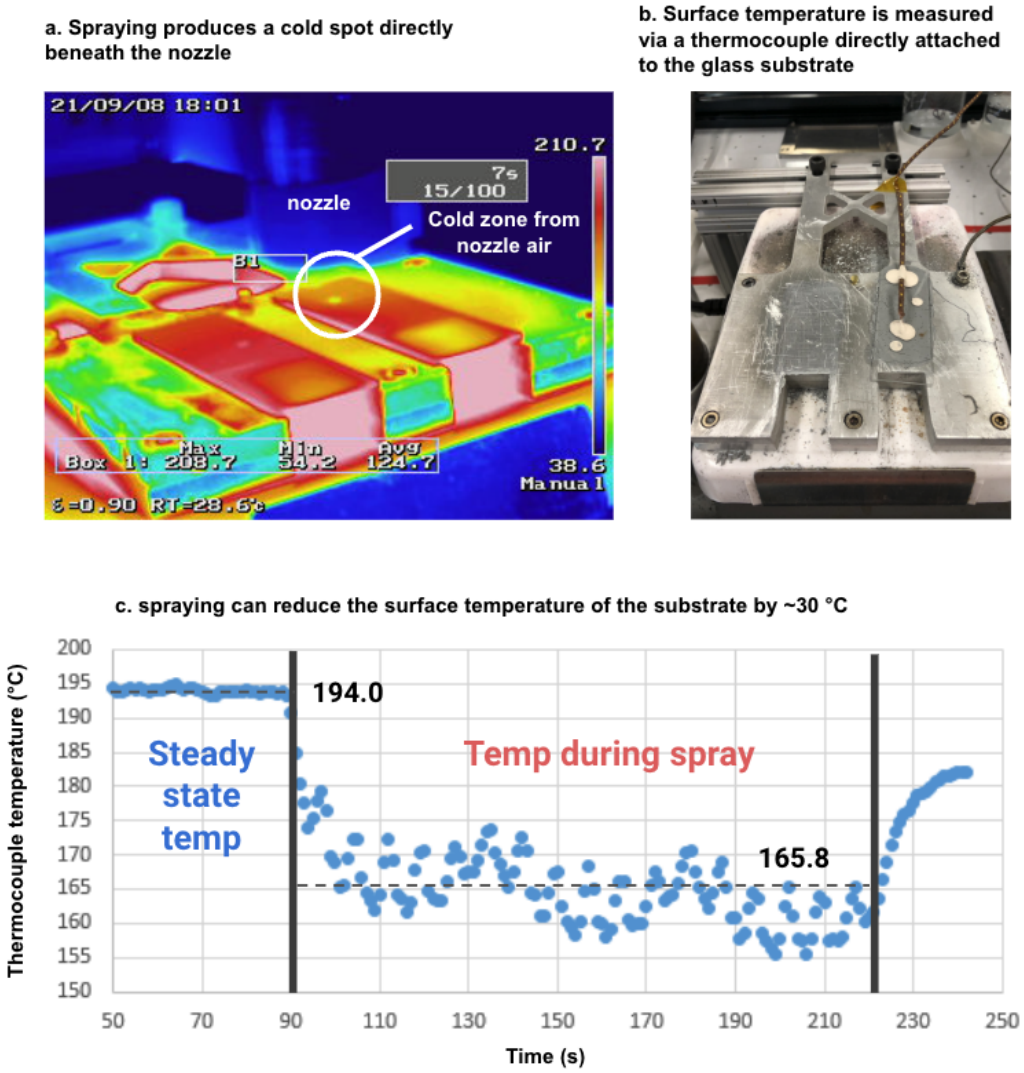

Figure S8 | Thermal camera and thermocouple data revealed that spraycoating can reduce the local surface temperature on the substrate by 30 °C. (a) thermal camera data showing the visible change in temperature on the surface of the substrate when spray-coating. **(b)** Using a thermocouple attached to a glass substrate with thermal cement, **(c)** the temperature of the surface of the substrate was measured to decrease by ~30°C during spray-coating.

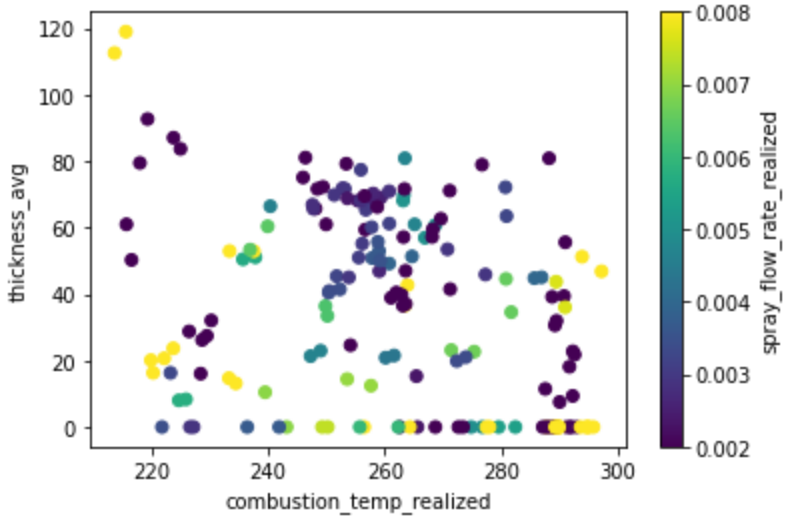

Figure S9 | Low ink flow rate is associated with higher film thickness at elevated temperature. At temperatures below 240°C, there is no clear trend between temperature, spray flow rate, and thickness. At temperatures above 240 °C, points with low flowrate (darker) have greater thickness. This suggests that the slower nozzle speed associated with lower flow rates is causing the nozzle to linger over the same spot for longer, cooling the sample more with the compressed air. This observation coincidentally lines up with the experiment performed in **Figure S7** where the two lines cross over each other at 240°C.

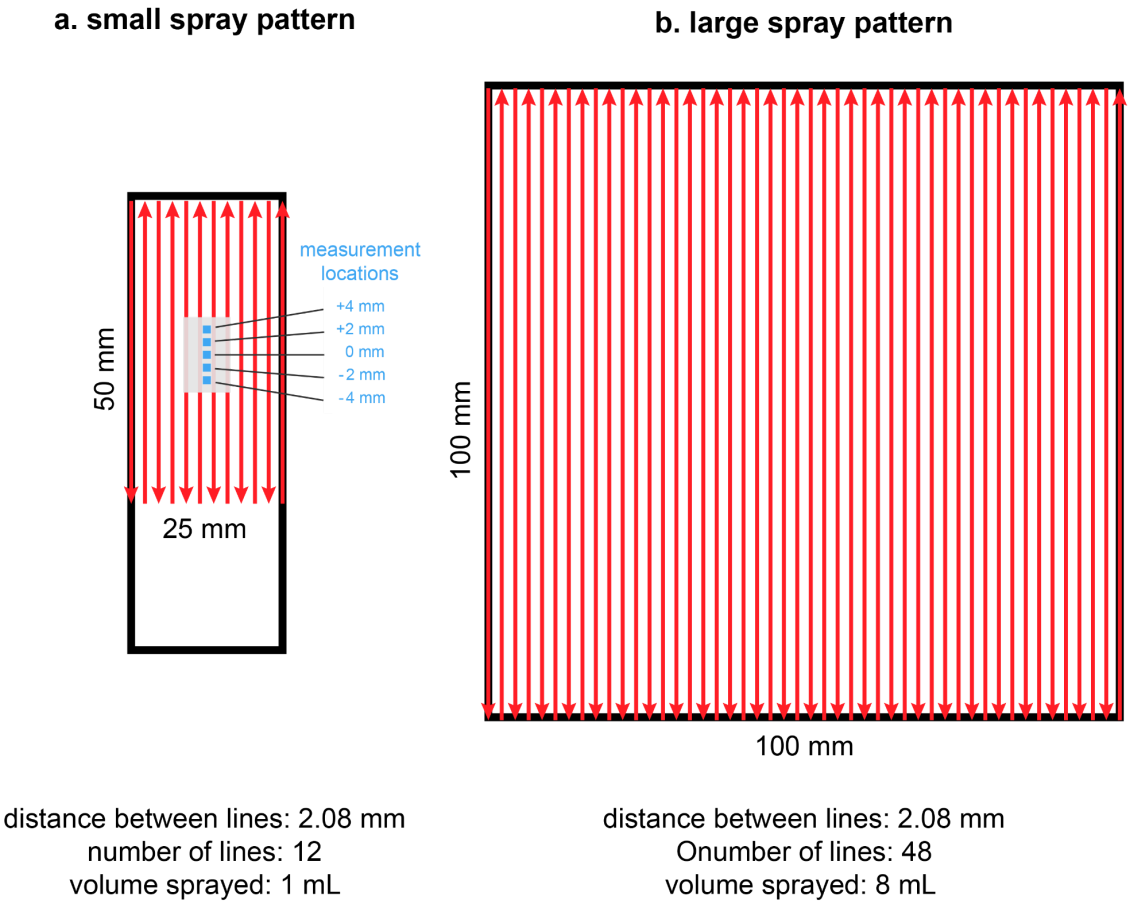

Figure S10 | The spray-coater pattern for (a) the small scale sample and (b) the large scale sample. The small spray pattern was used for all samples in the optimization campaign. Shown in blue are the locations for the 4-point probe and thickness measurements from the autonomous workflow (see supplementary methods - autonomous workflow steps 4-6). The center location (0 mm) is located directly in the center of the spray pattern.

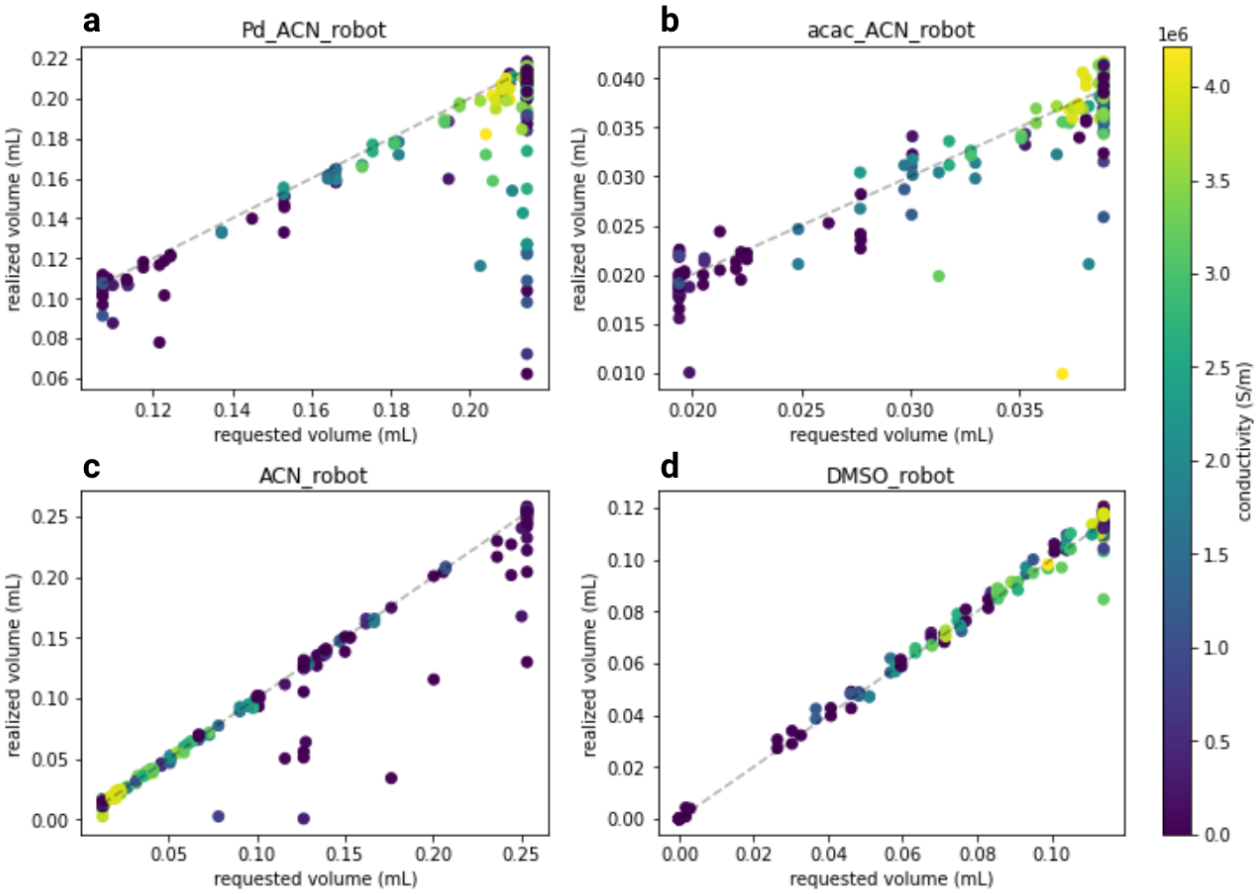

Figure S11 | Requested pipetting volumes versus realized pipetting volumes for each precursor ingredient for each sample in the optimization campaign. The requested volumes are based on experimental conditions requested for each sample. The realized volumes are based on the gravimetric feedback from the weighscale on the precursor mixing station. An error-free sample would fall on the grey dashed line in the center of each plot. Pipetting errors for palladium in acetonitrile (panel **a**) generally result in lower conductivity due to less Pd being in the precursor ink. More pipetting errors are reported for acetylacetone in acetonitrile (panel **b**) likely due to the small volumes that are requested.

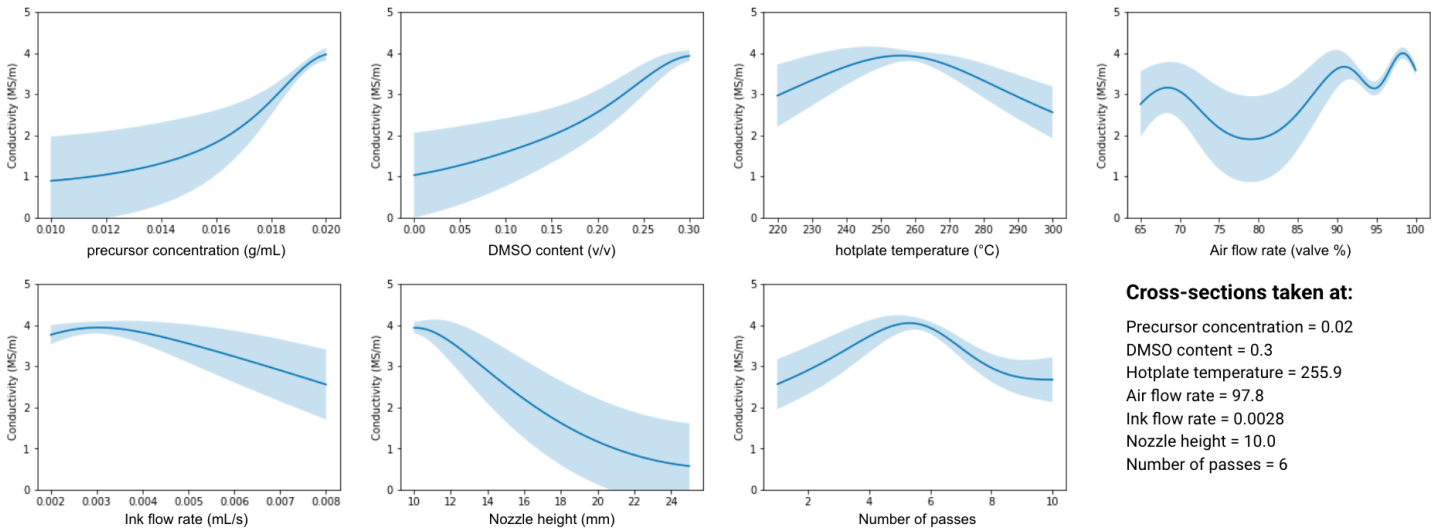

Figure S12 | Cross-sections of the final gaussian process model. These cross-sections were taken at the parameter values of the champion sample (sample 123 from experiment 61). Each plot was made by building the GP model over the dataset and predicting the conductivity at a fixed set of inputs (shown in the bottom right) and changing the only parameter of interest across its range. These plots represent what the optimizer thinks and how the conductivity changes as a function of each parameter.

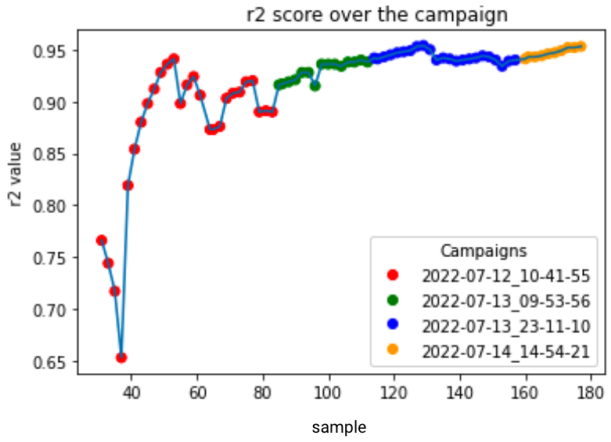

Figure S13 | The r^2 value of the model increases over the course of the campaign. The data in this plot represents the r^2 fit of the model each time a new experiment was requested during the optimization campaign (every other sample). There is no data before sample 30 because those samples were selected randomly (i.e. no model was built). The data is split by colour into four different campaign IDs from the four restarts of the self-driving lab.

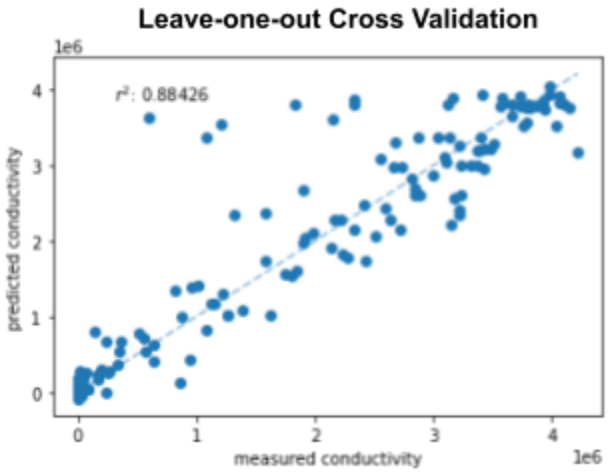

Figure S14 | The model was evaluated using leave-one-out cross validation (LOOCV) once the optimization was complete to assess the prediction accuracy of the model on unseen data, yielding a r^2 of 0.88. To perform LOOCV, one sample was removed from the dataset while the gaussian process model was built over the remaining dataset. The resulting model was used to predict the conductivity value of the sample that was left out. A higher r^2 score means that the model was better at predicting unseen data than a model with a lower r^2 score.

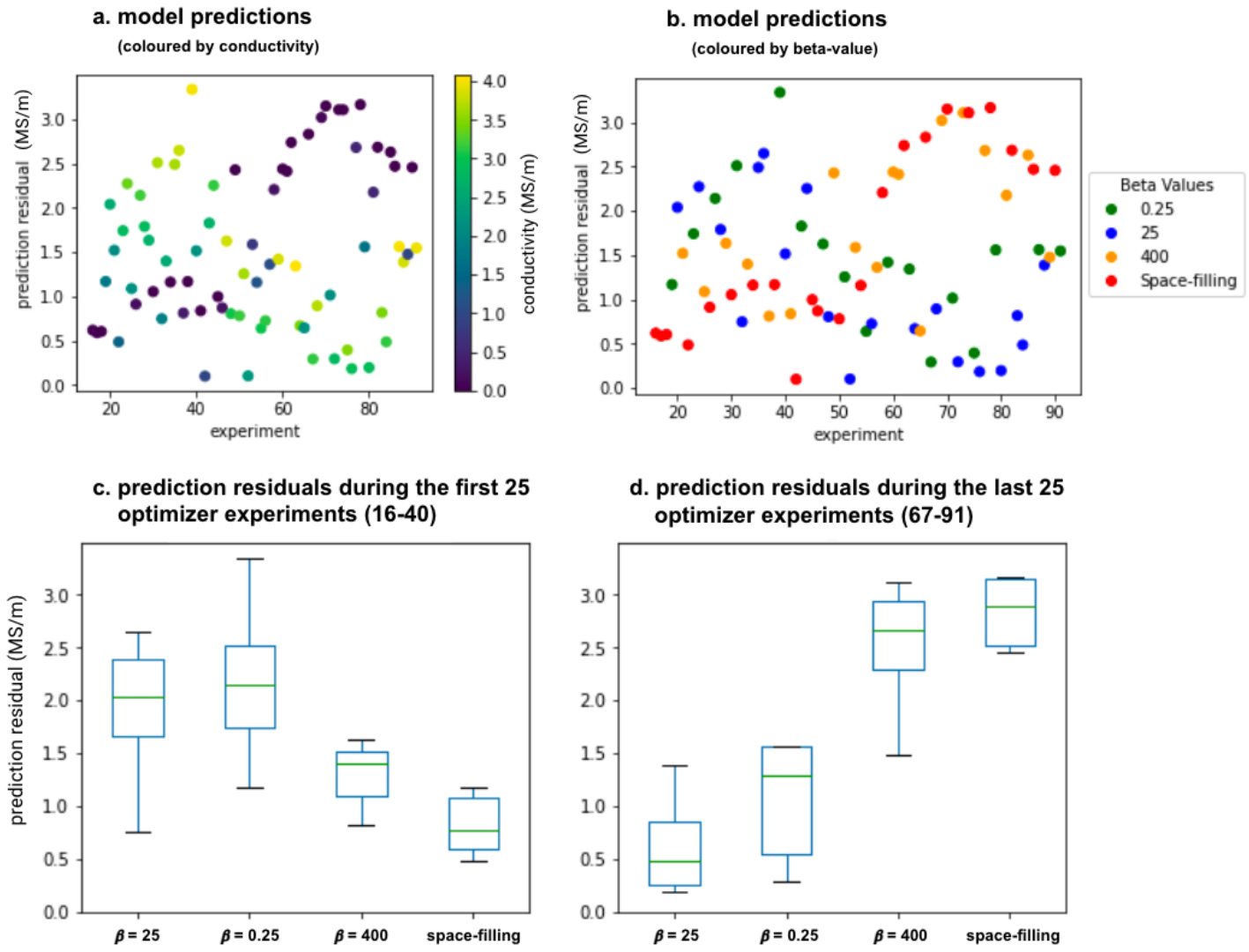

Figure S15 | Model predictions improved for low beta values (0.25 and 25) but not for the high beta value and space-filling. (**a**) The model improved at predicting where high conductivity samples lie but got worse at predicting where poor conductivity samples lie. (**b**) predictions of high conductivity are the result of the acquisition scheme of the optimizer. Exploitive low beta values (0.25 and 25) resulted in high conductivity points while exploratory high beta values (400) and space filling resulted in selection of experiments that were far from any previous data and generally resulted in poor conductivity and poor predictions of conductivity. (**c**) and (**d**) show that in the beginning of the campaign, the model was better at predicting exploratory points and by the end of the campaign it was better at predicting exploitive points. The data for these plots was generated by iteratively building a gaussian process model over the "finished" samples and predicting the conductivity of the next experiment. A lag of 4 data points was included to make the simulation similar to our parallelized robotic workflows. The residuals shown are the predicted conductivity value of the model compared to the measured value for the same experiment.

Supplementary references and acknowledgements

[1] B. P. MacLeod *et al.*, "A self-driving laboratory advances the Pareto front for material properties," *Nat. Commun.*, vol. 13, no. 1, pp. 1–10, Feb. 2022.
[2] "Luigi 2.8.13 documentation." https://luigi.readthedocs.io/en/stable/ (accessed Sep. 23, 2022).

We gratefully acknowledge the efforts of the contributors to the many open-source software packages used in this work:

- python (https://www.python.org/)
- NumPy (https://numpy.org/)
- pandas (https://pandas.pydata.org/)
- Matplotlib (https://matplotlib.org/)
- SciPy (https://www.scipy.org/)
- scikit-learn (https://scikit-learn.org/)
- Keras (https://keras.io/)
- TensorFlow (https://www.tensorflow.org/)
- Plotly (https://plotly.com/)
- LMFIT (https://lmfit.github.io/lmfit-py/)
- HyperSpy (https://hyperspy.org/)
- luigi (https://github.com/spotify/luigi)
- Ax (https://ax.dev/)
- BoTorch (https://botorch.org/)
- BoTorch License (https://github.com/pytorch/botorch/blob/main/LICENSE)
- PyTorch (https://pytorch.org/)
- Flask (https://flask.palletsprojects.com/)
- OBS Studio (https://obsproject.com/)
- Git (https://git-scm.com/)