

# Enhancing the Quality of MOF Thin Films for Device Integration Through Machine Learning: A Case Study on HKUST-1 SURMOF Optimization

Lena Pilz, Meike Koenig, Matthias Schwotzer, Hartmut Gliemann, Christof Wöll,\* and Manuel Tsotsalas\*

Metal–organic Frameworks (MOFs), especially as thin films, are increasingly recognized for their potential in device integration, notably in sensors and photo detectors. A critical factor in the performance of many MOF-based devices is the quality of the MOF interfaces. Achieving MOF thin films with smooth surfaces and low defect densities is essential. Given the extensive parameter space governing MOF thin film deposition, the use of machine learning (ML) methods to optimize deposition conditions is highly beneficial. Combined with robotic fabrication, ML can more effectively explore this space than traditional methods, simultaneously varying multiple parameters to improve optimization efficiency. Importantly, ML can provide deeper insights into the synthesis of MOF thin films, an essential area of research. This study focuses on refining an HKUST-1 SURMOF (surface-mounted MOF) to achieve minimal surface roughness and high crystallinity, including a quantitative analysis of the importance of the various synthesis parameters. Using the SyCoFinder ML technique, thin film surface quality is markedly enhanced in just three generations created by a genetic algorithm, covering 30 distinct parameter sets. This method greatly reduces the need for extensive experimentation. Moreover, the results enhance the understanding of the vast synthesis parameter space in HKUST-1 SURMOF growth and broaden the applications of MOF thin films in electronic and optoelectronic devices.

years ago by O. Yaghi and S. Kitagawa.<sup>[1–3]</sup> These materials, comprise two main components: metal ions or metal-oxo clusters acting as nodes, and organic linker molecules. The linkers are connected to the inorganic sub-building units through coordinative bonds. Nodes and linkers work in tandem resulting in the unique properties of MOFs, their choice allowing for precise control over the MOFs' geometric and chemical properties, such as pore size, aperture diameter, shape, reactivity for post-synthetic modifications, and interactions with other molecules (e.g., polarity, affinity to guest molecules).<sup>[4–8]</sup> This modular flexibility enables the optimization of MOFs for specific applications. Traditionally, MOFs are synthesized via solvothermal methods as powder bulk materials, with typical applications in gas storage, catalysis, and chemical sensing.<sup>[9]</sup>

However, bulk MOF materials, despite being easily produced in high yields (such as Basolith C300 from BASF), face significant limitations in many applications. A prominent example are optical applications, where the unwanted scattering

from randomly oriented MOF particles severely complicates the integration of this multi-functional material into devices.<sup>[10]</sup> To overcome these limitations, in 2007 Wöll and coworkers introduced a layer-by-layer (lbl) method to fabricate surface-anchored MOFs (SURMOFs).<sup>[11]</sup> SURMOFs can be grown on any substrate in an oriented manner using various layer-by-layer techniques, offering a new material platform for numerous applications.<sup>[12–14]</sup>

The introduction of SURMOFs and other types of MOF thin films provided the basis for integration into electronic devices, especially in the area of sensor technology.<sup>[15]</sup> As researchers strive to harness the unique attributes of SURMOFs, the challenges of establishing well-defined, sharp, and smooth interfaces and fine-tuning material properties have become a focal point: The integration of metal–organic frameworks for various applications into electronic devices has catalyzed the emergence of an entirely new research domain over the past decade. Heinke and coworkers publication in 2014 was a seminal work, exploring the mass transfer and surface barriers encountered when

## 1. Introduction

Metal–organic frameworks (MOFs) or Porous Coordination Polymers are highly porous and crystalline materials introduced  $\approx 25$

L. Pilz, M. Koenig, M. Schwotzer, H. Gliemann, C. Wöll, M. Tsotsalas  
Institute of Functional Interfaces  
Karlsruhe Institute of Technology  
Hermann-von-Helmholtz-Platz 1, 76344 Eggenstein-Leopoldshafen,  
Germany  
E-mail: [christof.woell@kit.edu](mailto:christof.woell@kit.edu); [manuel.tsotsalas@kit.edu](mailto:manuel.tsotsalas@kit.edu)

The ORCID identification number(s) for the author(s) of this article can be found under <https://doi.org/10.1002/adfm.202404631>

© 2024 The Author(s). Advanced Functional Materials published by Wiley-VCH GmbH. This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial](https://creativecommons.org/licenses/by-nc/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

DOI: 10.1002/adfm.202404631

incorporating guest molecules into MOFs, a crucial step toward the utilization of MOFs in QCM devices.<sup>[16]</sup> More recently, Mohan et al. have made significant strides, showcasing the successful integration of MOF luminescence and electrochemical sensors in devices tailored for cancer biomarker detection.<sup>[17]</sup> Ohira et al. contributed to the field with their 2015 paper on a novel fiber optic sensor, employing a metal–organic framework coating as the main innovation.<sup>[18]</sup> In a comprehensive review in 2017, Stassen et al. delved into the multifaceted applications of MOFs in electronics and chemical sensing, covering their use in digital circuits, field-effect transistors, and mass-sensitive sensors, among other technologies.<sup>[19]</sup> The rapid progression in MOF-based electronic device research is evident through the sheer volume of papers published annually, with frequent updates in reviews, underscoring the substantial interest and impact in this specialized application area of SURMOFs.<sup>[20–25]</sup> Another illustrative example of the impact of SURMOF roughness on the physical properties of a thin film is demonstrated in the work of Zhi-Gang Gu et al.<sup>[26]</sup> In their study, they showed that by selecting specific experimental synthesis conditions, the roughness of a SURMOF could be significantly reduced by employing ultrasonication, leading to a notable enhancement in the optical quality of the MOF thin film. Also in cases where MOF-hetero-interfaces are important, the quality of the interface between two different MOF thin films can be crucial, e.g. in the case of photon upconversion.<sup>[27]</sup>

Bridging these discussions, it becomes clear that while the integration of MOFs into electronic devices has made remarkable strides, a persistent challenge remains in achieving smooth surfaces during their manufacturing. Traditional methods like spray synthesis, spin-coating, and vapor assisted conversion (VAC) often fall short in this regard. The importance of smooth surfaces cannot be overstated; they are crucial for maintaining structural integrity and facilitating effective host-guest interactions, which are key to the functionality of MOFs in various applications.

The initial hurdle for achieving a smooth SURMOF surface is starting the growth of the first MOF layers on a functionalized substrate. In the subsequent steps of the lbl process, the challenge lies in achieving smooth outer surfaces to fully harness the distinctive attributes of MOF materials. A further overarching challenge involves the optimization of desired properties within the MOF system, such as achieving excellent crystallinity and low defect densities.<sup>[28]</sup> There are examples where the defect densities in SURMOFs are substantially smaller than in the corresponding bulk materials.<sup>[29]</sup>

In view of the numerous parameters governing the lbl growth, the corresponding parameter space is extensive, and optimizing properties such as structural quality and surface roughness represents a time-consuming effort. Since the low-temperature lbl growth is governed mostly by kinetic rather than thermodynamic control, optimization of growth parameters is typically based on a trial-and-error approach. Experience has demonstrated that often simultaneous variation of multiple parameters is required, thus requiring a departure from classical optimization techniques, which involve altering one parameter at a time to analyze the resulting effects. Previous studies have already demonstrated the effectiveness of machine learning techniques as valuable tools for optimizing and discovering functional materials in various contexts.<sup>[30–32]</sup> Building on this foundation, this intricate scenario

calls for employing machine learning techniques to navigate this complexity effectively, as illustrated in **Figure 1**, ultimately enabling the optimization of properties with minimal experimental effort.

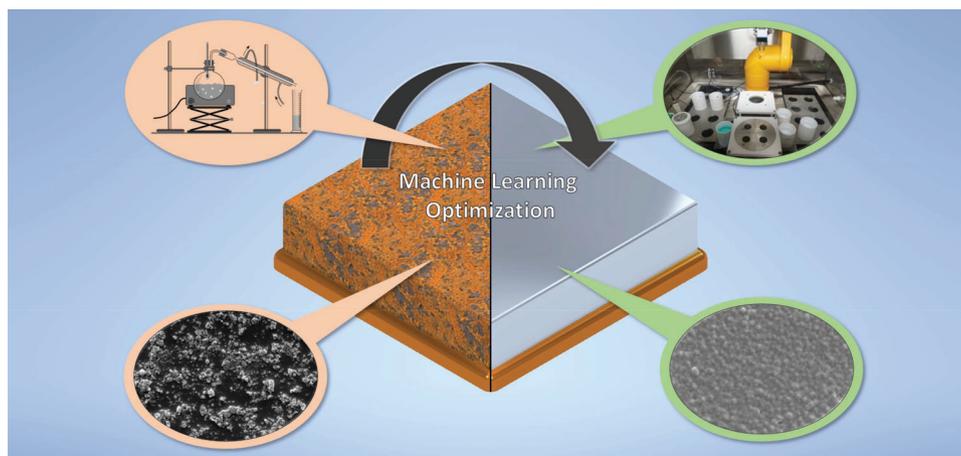
In addition to achieving a particular property of a SURMOF, the understanding of the system's growth mechanism is equally vital. Integrating ML methods in the SURMOF synthesis can aid in unraveling these mechanisms, potentially revealing fundamental principles.

In this work, we utilized ML techniques to simultaneously enhance the crystallinity and reduce the surface roughness, that ML methods are well suited to steer the growth direction of HKUST-1 MOF thin films in lbl deposition by optimizing various deposition parameters.<sup>[34,35]</sup> Encouraged by the rather impressive success of this previous work, we conducted an in-depth investigation focusing primarily on reducing the roughness of MOF thin films - a critical attribute for several MOF-based applications, as described in the discussion above. Besides, we carried out a comprehensive exploration of the growth mechanism, particularly examining the impact of specific parameters during the early stages of the lbl process. This optimization required the simultaneous adjustment of seven different parameters, thus going substantially beyond the previous study where only five parameters were considered.

In general, ML methods require high-quality of experimental results, unwanted variations e.g., resulting from poor control of synthesis parameters can substantially affect the success of the optimization. To ensure the reproducibility of synthesis procedures, an industrial synthesis robot was employed.<sup>[26,34]</sup> All syntheses were conducted in a glove box under inert conditions. In all cases, a 40-cycle layer-by-layer deposition process was employed.<sup>[36]</sup> The samples were then subjected to X-Ray diffraction (XRD) analysis to assess phase purity and crystallinity, and to ellipsometry for a quantitative determination of surface roughness. All data is made publicly available in the Chemotion repository, therefore supporting sustainable scientific research and facilitating future discoveries.<sup>[37]</sup>

As previously mentioned, the objective was not solely to attain a smooth SURMOF surface but also to probe the influences in the growth mechanism. Consequently, seven parameters were selected: the concentration of metal salt and organic linker, the amount of modulator, ultrasonication, and spray cleaning duration, along with two additional variables set for the metal salt and organic linker concentration during the first three synthesis cycles. The underlying hypothesis posited that the concentrations of the metal salt, organic linker, or both during the initial cycles, while establishing the boundary layer on the substrate's surface previously covered only by a self-assembled monolayer (SAM), would be the pivotal factors influencing SURMOF growth and for the eventual attainment of a smooth surface.

The SyCoFinder operates through three sequential steps, starting with the creation of its own training dataset from practical experiments. This initial dataset then serves as the basis for the subsequent optimization process, which utilizes a genetic algorithm. A more detailed description of this process can be found in the experimental section (See Machine Learning Method). As third and last step of the SyCoFinder an evaluation of the chosen variables and their respective importance can be carried out,



**Figure 1.** Exemplary illustration of advancement when using machine learning methods and automated systems; On the left side a non-optimized rough and defective MOF is symbolized and on the right side the machine learning optimized smooth and even MOF thin film is depicted.

potentially offering valuable insights into their impact on the defined quality criteria.

In previous studies, this investigation concerning the influence of the chosen parameters has led to interesting insights in the growth mechanism of HKUST-1, for instance revealing the importance of water on the formation of a certain orientation and the utilization of linker in excess to the metal salt.<sup>[34]</sup> Therefore, with the ambition of investigating the impact of the initial cycles on the surface smoothness, this final step was deemed especially important.

## 2. Results and Discussion

### 2.1. Evolution

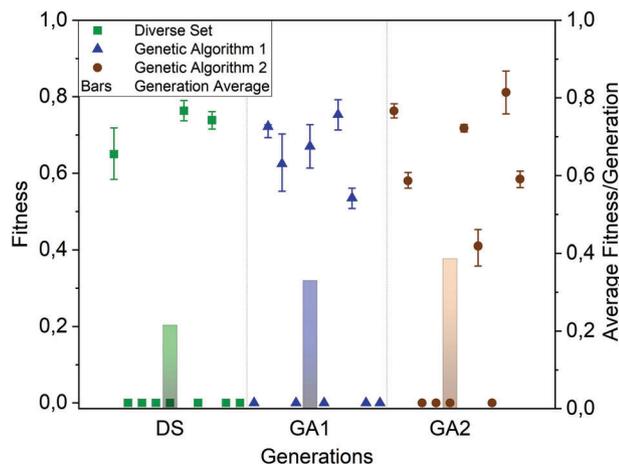
The overarching aim was to create a high-quality HKUST-1-SURMOF surface, characterized by low roughness and high crystallinity. Seven synthesis parameters, detailed in **Table 1**, were selected for variation. The ranges for variations were determined from experimental constraints, including solubilities, solvent consumption during spray cleaning, and inherent limitations from prolonged ultrasonication. Additionally, results from previous work served as a guide for defining the range in the case of the growth modulator.<sup>[38,39]</sup> It is important to note that, except for the amounts of metal salts and linker, the potential for reaching a zero value was always considered. Each solution had a volume

of 210 mL. The parameter sets for all generations containing the actually applied values for each variable are to be found in **Tables S1– S3** (Supporting Information).

As previously mentioned, the metal and linker concentrations during the first three cycles (metal salt 2 and linker 2) were optimized independently from the concentration during steps 4 to 40. Following immediately on these initial cycles, the metal and linker solutions were adjusted for cycles 4 through 40 in every experiment. Notably, the remaining variables were held constant throughout their respective experiments, and a consistent 40-cycle count was maintained across all experiments. The analysis of all samples was subsequently conducted using X-Ray diffraction to determine the phase identity and crystallinity and ellipsometry to analyze the surface roughness. Phase identity is an exclusion criterion, thus simply evaluating the formation of pure HKUST-1 in principal. Crystallinity represents the quality of the respective SURMOF in terms of XRD. The roughness however, inherits both quality and exclusion criteria. The roughness itself is the quality indicator emerging from ellipsometry. When performing evaluations in ellipsometry, a mathematical model is developed and adjusted to match the measurement curve. For comparability, essential for ML optimizations, two threshold values are set: the MSE (Mean Squared Error) (more detailed explanation is to be found in Fitness) and a roughness-to-thickness ratio cut-off. These thresholds serve as exclusion criteria in the roughness assessment. Details on calculating each term are to be

**Table 1.** The seven parameters chosen to be varied and their according ranges.

Variable	Range	Applied on cycle number
Metal salt	0,02–6,00 mmol (0,1 - 28,6 mmol/L)	4-40
Linker	0,02–10,00 mmol (0,1 - 47,6 mmol/L)	4-40
Amount of water (modulator)	0,0–40,0 mL	1-40
Cleaning time via ultrasonication	0–100 s	1-40
Cleaning time via spray-unit	0–5 s	1-40
Metal salt 2	0,02–6,00 mmol (0,1 - 28,6 mmol/L)	1-3
Linker 2	0,02–10,00 mmol (0,1 - 47,6 mmol/L)	1-3



**Figure 2.** Fitness value development of the machine learning optimization over three generations following the color code: Diverse Set (green squares), first genetic algorithm, (GA1) (blue triangles) and second genetic algorithm (GA2) (brown dots) and average bars in the according colors.

found in Fitness. The resulting fitness value ranging from zero to one, then ranks the experiments according to Equation (1), allowing the genetic algorithm to optimize.

$$\text{fitness} = \text{fitness}(\text{phase identity}) * \text{fitness}(\text{crystallinity}) * \text{fitness}(\text{roughness}) \quad (1)$$

**Figure 2** illustrates the progression of the overall fitness calculated for each experiment using Equation (1). Notably, with each generation more experiments receive a fitness value, thus indicating that more experiments fulfill the required criteria, to receive a fitness value.

To comprehend this development, it is instructive to examine the individual evaluation developments of crystallinity and roughness terms contributing to the overall fitness calculation. Initially, the low number of successful ratings in early evaluations (Diverse Set) is attributed to stringent experiment exclusion criteria. The fitness evaluation was initially designed to be very restrictive, thus limiting the acceptance of experiments for evaluation. While all samples from the first generation showed acceptable to excellent crystallinity (see Figure ?), many were disqualified due to shortcomings in other criteria, particularly roughness determined by ellipsometry. This rigorous approach was crucial to ensure the production of reliable data. The methodology for including or excluding experiments is delineated in Equation (6). Furthermore, all values contributing to the final fitness value of an experiment are to be found in the Supporting Information in Tables S4–S6. **Figure 3b** displays the fitness of roughness for experiments falling within the evaluation range, while **Figure 3a** showcases their corresponding crystallinity fitness values. The corresponding X-Ray diffractograms are also summarized in the Supporting Information (See Figures S2–S7, Supporting Information).

After three generations, the optimization yielded a parameter set reaching an 81% fitness, corresponding to 92% crystallinity and a roughness value of 6.52 nm (an 89% roughness fit). The

achievement of such a high fitness level with seven variables in just ten experiments per generation is particularly remarkable due to the exponential increase in complexity with each additional variable. In previous studies, with five variables, there were effectively two variations per variable per generation, providing a more manageable exploration scope.<sup>[31,34]</sup> However, incorporating seven variables intensifies the challenge, as the parameter space grows exponentially, making each decision point more critical and the successful navigation through this vast space a notable accomplishment. This exponential growth means that every new variable dramatically expands the potential combinations, making it increasingly difficult to pinpoint the optimal settings within the limited number of experiments.

The SEM picture in **Figure 4a** shows a successful sample from the Diverse Set (08-DS), **Figure 4b** the best one in GA2 (09-GA2) underlining its exceptional smoothness and therefore success of the optimization.

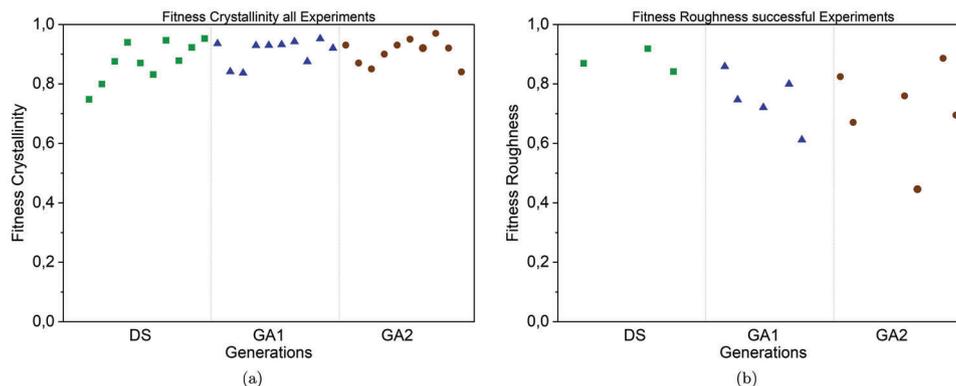
Although the fitness difference between the best experiment in the Diverse Set (76%) and the optimal GA2 experiment (81%) might appear relatively modest, it is crucial to note the broader objective: The best Diverse Set experiment (06-DS) resulted in a lower roughness of 4.54 nm, surpassing the GA2's best (09-GA2). However, in this case, the MOF thin film was rather amorphous, the intensity of the diffraction peaks was low. Therefore, this parameter set did not achieve the goal of simultaneously reaching low roughness and high crystallinity. The best parameter set attained a crystallinity of 91.6% alongside a roughness of 6.52 nm, outperforming the diverse set's best experiment which, despite its superior roughness of 4.54 nm, only reached a crystallinity of 83.2%. This highlights the success of balancing both goals in the optimization process. The achievement of a smooth and crystalline SURMOF is noteworthy, especially considering the immense parameter space created from seven variables. With only 10 experiments per generation, this allows for fewer than two variations per variable in each generation. Consequently, the significant improvement in thin film quality within just 30 experiments is remarkable.

Moreover, to validate the hypothesis concerning the impact of the initial three cycles on SURMOF growth and surface roughness, more than three data points from the Diverse Set were needed for a comprehensive assessment. This necessity for further data points represented another reason to continue with the optimization process rather than ending it prematurely after the Diverse Set.

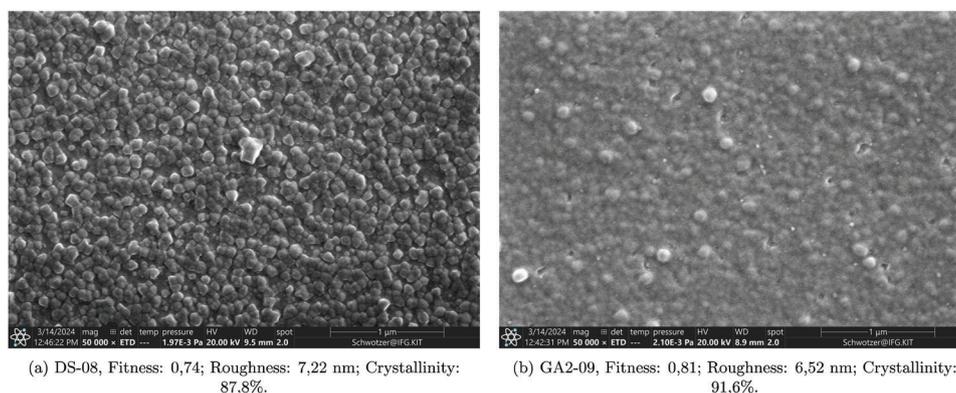
## 2.2. Evaluation of Relative Importance of Variables

An examination of **Figure 5**, depicting the relative importance of variables, provides valuable insights in this endeavor. In this final step, the SyCoFinder was provided with all data from the optimization steps (GA1 and GA2) to assess the most significant variables with respect to user-defined objectives. The plot showcases the relative importance of these variables, ranging from 0 to 1.

The pivotal factor in this assessment is the concentration of the metal salt utilized in the first three cycles, followed closely by the linker concentration in the first three cycles, substantiating the hypothesis regarding their influence on the synthesis outcome.

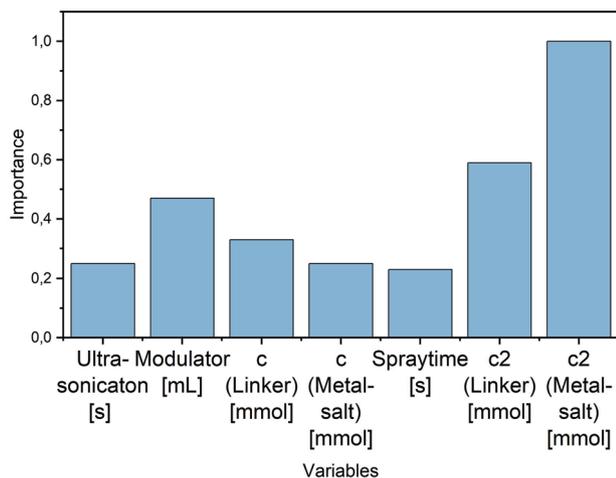


**Figure 3.** Color code for both figures: Diverse Set (green squares), first genetic algorithm (GA1) (blue triangles) and second genetic algorithm (GA2) (brown dots). a) Fitness of crystallinity of all experiments, b) Fitness of roughness of all successful experiments.



**Figure 4.** SEM pictures of a) a sample from the Diverse Set, yielding 74% fitness and b) the best sample from the second genetic algorithm (GA2) yielding a fitness of 81%, showing a very uniform and smooth film.

To gain deeper insight, **Table 2** is considered, presenting all successful parameter combinations of GA1 and GA2 arranged by decreasing fitness. The original parameter sets can be found in the Supporting Information in Tables **S1–S3**.



**Figure 5.** Relative importance of all experiments from GA1 and GA2 for the selected variables (c-linker, c-metal, modulator, ultrasonication time, spray cleaning time, c2-linker, c2-metal) on a scale of 0 to 1.

Considering the best three experiments, a clear trend is visible: The initial cycles show higher concentrations of both metal salt and linker compared to the later cycles. This leads us to the assumption that increased concentrations early in the process help achieve a more complete coverage of the surface with base layers, thereby reducing the likelihood of island formation. By establishing a well-covered base framework in the early cycles, the potential for a defect-free SURMOF structure is significantly enhanced, setting a primary structure that guides the orderly growth of subsequent layers. At this stage, molecular interactions and lattice formation are already underway, allowing the crystal growth to proceed under less intense conditions. Consequently, lower concentrations in later cycles are generally sufficient, as the focus shifts from initiating to maintaining and expanding the crystal structure. Yet, both the initial cycles (1–3) and the later cycles (4–40), the concentration of the linker is notably higher than the metal salt's concentration.

The third most important variable is the amount of modulator in the system, aligning with existing knowledge that water significantly influences the crystallization of SURMOFs. However, water also strongly influences the orientation of SURMOFs, which we did not consider in the fitness function in this study.<sup>[34,38]</sup> Even though water amounts on the upper limit of 30–40 mL seem less favorable, further assumptions about its importance are limited.

**Table 2.** All successful parameter sets ordered by decreasing fitness.

Fitness	Ultra-sonication [s]	Modulator (Water) [mL]	c (Linker) [mmol L <sup>-1</sup> ] Cycle 4–40	c (Metal) [mmol L <sup>-1</sup> ] Cycle 4–40	Spray-cleaning [s]	c2 (Linker) [mmol L <sup>-1</sup> ] Cycle 1-3	c2 (Metal) [mmol/L] Cycle 1-3
0,81	36	7,81	39,48	17,86	3	48,67	20,38
0,76	45	7,75	36,86	6,71	0	55,43	26,00
0,75	0	18,15	33,24	17,90	3	45,52	30,38
0,72	0	0	47,62	0,10	5	34,95	28,57
0,72	0	24,3	34,19	0,14	4	36,67	5,48
0,67	12	2,93	47,62	0,14	5	31,67	28,57
0,62	81	0	12,00	28,57	5	37,76	14,14
0,58	58	2,79	22,71	28,57	5	39,14	4,95
0,58	10	7,24	12,95	5,10	5	40,05	9,29
0,54	63	22,61	0,24	0,24	2	1,57	7,57
0,41	5	5,26	47,19	12,24	4	42,90	30,19

The variables for the cleaning methods employing ultrasonication and spray cleaning are located at comparatively low values, which is substantiated by their values noted in Table 2 not really depicting any notable trends.

It is challenging to draw definitive conclusions about the direct impact of specific parameters on achieving low roughness and high crystallinity, as clear trends are primarily visible in the top three experiments only. Consequently, this evaluation of variable importance and the conclusions derived should be viewed as logical deductions rather than concrete evidence. Nevertheless, these findings serve as a solid basis for more detailed investigations, providing valuable insights into the significant factors influencing the outcomes.

Additionally, based on this study, an existing functionality of the SyCoFinder could prove useful in accelerating further optimizations in similar systems. When creating the Diverse Set, variables can be weighted. Leveraging the insights from our variable importance evaluation, strategic weighting can introduce a beneficial bias, accelerating optimizations for systems with similar chemistry to those studied.

However, the hypothesis concerning the substantial impact of the first three cycles was exclusively related to the concentration variables, being emphasized through the evaluation of the importance of variables. This observation underscores the significance of meticulous control and optimization of these early cycles for achieving the desired material properties. However, future studies could delve further into exploring the roles of other parameters during these initial cycles, such as variations in cleaning methods, to determine their potential influence on the synthesis outcome.

### 3. Conclusion

In this study, we employed a structured approach to analyze the impact of distinct variables controlling the layer-by-layer synthesis on the quality of MOF thin films, focusing on HKUST-1 SURMOFs and particularly emphasizing the initial synthesis cycles. Our investigation highlighted the critical importance of the first three cycles, especially the concentrations of metal salt and linker. We discovered that the parameters set in these early cycles play a crucial role in determining the final characteristics of the de-

posited MOF thin films, especially in terms of crystallinity and surface smoothness.

Our findings underscore the necessity of precise control and optimization of these initial cycles, specifically in terms of concentration parameters, to obtain the desired material properties. This detailed understanding paves the way to enhancing synthesis processes in other SURMOF systems by applying these insights.

A highlight of our research was identifying a highly effective synthesis method, achieving an impressive 81% fitness level. This method produced a material with exceptional crystallinity, quantified at 91%, and a surface roughness of just 6.52 nm, translating to an 89% roughness fit. Remarkably, this was attained by adjusting seven different parameters across only three ML generations. However, we recognize that further exploration could potentially reveal even more optimal methods within the extensive parameter space.

The results of this study not only contribute significantly to the field of MOF thin films and their integration into devices but also establish a foundation for future research. The methodologies and insights gleaned here will guide the optimization of synthesis techniques to develop MOF thin films and heterolayers with superior qualities, opening new avenues for the application of MOFs in various technological domains.

### 4. Experimental Section

**Materials:** For the automated synthesis of HKUST-1 as a SURMOF gold-coated silicon wafer pieces of  $1 \times 3 \text{ cm}^2$  in area were used as a substrate. The uncoated wafers were purchased from Siegert Wafer and pre-cut on the reverse side by MaTeck to ensure consistent piece sizes.

For synthesis and cleaning steps ethanol (abs.) (AnalaR NORMAPUR) was used as a solvent and purchased from VWR, copper-di-acetate was purchased from ACROS ORGANICS and trimesic acid (benzene-1,3,5-tricarboxylic acid) from Sigma–Aldrich.

**Synthesis:** Prior to the synthesis the gold-coated silicon wafer pieces were immersed into a SAM-solution of 1, 4 mg 16-Mercaptohexadecanoic acid in 25, 0 mL acetic acid (glacial) and 225 mL ethanol (abs.) for 72 h.

In the following, these pre-functionalized substrates were mounted into a sample holder and inserted into the robotic set-up (See S1, Supporting Information and Ref. [34]). Housed within a glovebox, the setup maintains inert under reproducible synthesis conditions. The industrial six-axis robot

from Stäubli then immerses the substrates subsequently into an ethanolic solution of copper-di-acetate, followed by one or more cleaning steps in pure ethanol, then into a linker solution of trimesic acid (BTC) in ethanol and water as a modulator, again followed by one or more cleaning steps. These cleaning steps were either simply dip-rinsing, meaning immersing the substrate into another vessel containing ethanol or spray-cleaning, where the substrates were rinsed from both sides with ethanol dispersed from a nozzle with a certain pressure or ultra-sonic-cleaning where the substrates were also immersed into a vessel containing ethanol but with simultaneous ultrasonic treatment. This sequence represents one cycle. The respective robotic sequences for each parameter-combination is to be found in the Supporting Information (See “Possible Sequences” in S4, Supporting Information). Each experiment is carried out for 40 cycles, with certain concentrations of metal-salt and linker solution for the first three cycles and different concentrations for the following 37 cycles.

The dwelling times in each vessel as well as the concentrations of linker and metal-salt solutions and the corresponding combination of cleaning steps depends on the parameter combination emerging from the machine learning method. For example if the ultrasonication time was set to zero, the ultrasonic cleaning step was replaced by dip-cleaning. A list of the possible combinations or robotic programs respectively as well as the parameter sets provided by the SyCoFinder were to be found in the Supporting Information (see Tables S1–S3 as well as “Possible Sequences” 4, Supporting Information).

**Machine Learning Method:** The optimization process for achieving pure and highly crystalline HKUST-1 thin films with very low roughness followed the procedure outlined by the SyCoFinder web application. The SyCoFinder process commences with what was referred to as a “Diverse Set,” where the initial dataset for the genetic algorithm to be trained on was generated, rendering prior data unnecessary. During this step, the variables were defined, their ranges established, and the number of experiments determined. From this, the Diverse Set was created, comprising ten sets of parameter combinations each representing one experiment.

Based on these parameter sets, the synthesis of HKUST-1-thin films was conducted, and their characteristics were analyzed to determine their fitness. The web application then utilized the same parameter sets, along with their corresponding fitness values, to generate a new set of parameter combinations for the subsequent generation. For this process, a genetic algorithm was employed, which develops suggestions for more optimized parameter sets based on evolutionary principles. Specifically, this involved recombining parent parameter sets and applying a mutation factor to generate further optimized child parameter sets. For detailed information on the exact procedure, the reader is referred to the work of Moosavi et al.<sup>[35]</sup> This process of evolution through generations was repeated twice, resulting in a total of 30 experiments, until satisfactory results were achieved.

The optimization process employed a genetic algorithm after the initial Diverse Set was established. This algorithm involved recombination and mutation of the ranked parameter sets, with the ultimate goal of obtaining smooth, pure and highly crystalline HKUST-1-thin films.

**X-Ray Diffraction:** All diffractograms were captured using a Bruker D8 Advance diffractometer in  $\theta$ – $\theta$  geometry, ranging from 3 to 20  $2\theta^\circ$ . The instrument was equipped with a LYNXEYE position-sensitive detector featuring 192 active stripes. Additionally, an extended range from  $2\theta^\circ = 37$  to 40 was specifically recorded to detect the characteristic substrate gold diffraction peak, serving as a reference point. Each measurement was taken at a 2-s interval.

To ensure accuracy, each measurement underwent height error correction and background correction using the DIFFRAC.EVA software version 5.2.0.3 by Bruker AXS. The assessment of crystallinity utilized a predefined routine by Bruker, while the phase identity was conducted by comparing the diffractograms to a simulated HKUST-1 powder diffractogram.

**Ellipsometry:** Spectroscopic ellipsometry measurements were conducted using an M2000 instrument from J.A. Woollam Co. Inc., based in Lincoln, NE, USA. The measurements were carried out at an angle of incidence of  $70^\circ$  within the spectral range of 370–1000 nm under environmental conditions. To analyze the experimental data of the SURMOF thin film, a Cauchy-fit model was employed, utilizing the instrument’s software CompleteEase (V5.19). This model included roughness as an additional

effective medium layer, consisting of a 50–50 mixture of the ambient with the layer material. The optical characteristics of the gold-coated silicon substrate were determined through a reference measurement of the uncoated substrate performed prior to the synthesis.

**Fitness:** The *fitness*, represented as a value ranging from 0 to 1 and serving as a metric to evaluate each experiment, was determined by

$$fitness = fitness(phase\ identity) * fitness(crystallinity) * fitness(roughness) \quad (2)$$

Each of the multiplied terms refers to a different characterization technique, each referring to the average of measurements from two replicated samples. *phase identity* as well as *crystallinity* emerge from X-Ray diffraction measurements, while the term for *roughness* emerges from ellipsometric measurements.

The *fitness(phase identity)* expresses a simple exclusion criterion by differing into two cases:

$$fitness(phase\ identity) = \begin{cases} 1 & \text{if } \frac{f_1+f_2}{2} = 1 \\ 0 & \text{else} \end{cases} \quad (3)$$

In the first scenario, the term *fitness(phase identity)* yielded a value of one when the position of signals in the X-Ray diffractogram from the observed sample aligns with those in a simulated counterpart. Conversely, in the second scenario, this term resulted in a value of zero, indicating a lack of correspondence, which may manifest as the absence of clear phases or the complete absence of signals. The average value of the two diffractograms, denoted as  $f_1$  and  $f_2$ , which could each take on either the value 1 for a positive match or 0 in the case of amorphous samples, was calculated from two samples per experiment. This consistent practice of measuring two samples serves as a means to ensure the reproducibility of the results.

As previously stated, the calculation of *fitness(crystallinity)* was conducted using a built-in routine of the Bruker Software “Diffrac.eva”, which evaluates *crystallinity* as a percentage value. This assessment involves the following equations: the global area was defined as the integral under the entire, uncorrected measurement curve, while the reduced area is the integral under the curve after excluding the amorphous proportion. An example of this fitting process employing a built-in routine from Bruker is provided in the Supporting Information (See Figure S9, Supporting Information).

$$fitness(crystallinity) = 100 - \% \text{ Amorphous} \quad (4)$$

$$\text{Amorphous}[\%] = \frac{\text{Global Area} - \text{Reduced Area}}{\text{Global Area}} * 100 \quad (5)$$

The *fitness(roughness)* was derived from ellipsometric measurements, where the primary objective was to match data generated from an optical model to closely align with the observed data. A visual example of this fitting process is provided in the Supporting Information (See Figure S10, Supporting Information). The parameters *roughness* ( $R$ ) and *thickness* ( $T$ ), measured in nanometers (nm), were directly obtained from this analytical process. To prioritize low roughness relative to thickness, the ratio of these two values was subtracted from one, thereby aligning lower roughness values with a *fitness(roughness)* closer to one. This approach ensures that roughness and thickness were proportionally related, acknowledging their inherent correlation.

$$fitness(roughness) = \left(1 - \frac{roughness}{thickness}\right) * mse * R/T\text{-ratio} \quad (6)$$

$$mse = \begin{cases} 1 & \text{for } MSE < 20 \\ 0 & \text{for } MSE > 20 \end{cases}$$

$$R/T\text{-ratio} = \begin{cases} 1 & \text{for } \frac{R}{T} * 100 < 60\% \\ 0 & \text{for } \frac{R}{T} * 100 > 60\% \end{cases}$$

To maintain consistency and comparability across all measurements, the same analysis procedure was applied uniformly. However, this necessitated the establishment of a threshold to discard unreliable results. One such metric was the dimensionless Mean Squared Error (MSE), a measure for the match between the generated and the measured data and provided by the evaluation software during curve fitting. The higher this value, the more the experimental surface differed from the ideal layer model assuming sharp interfaces with minimal roughness. Based on experience from previous studies, measurements with a MSE above 20 were deemed unreliable. Consequently, the conditional logic of the second term excluded measurements with an MSE exceeding this threshold. Additionally, the roughness-to-thickness ratio, based on empirical observations, was set with a cut-off value of 60%. This threshold was designed to allow the machine learning algorithm to learn from sub-optimal results without being misled by them.

## Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.

## Acknowledgements

The authors acknowledge the use of OpenAI's ChatGPT for assistance in refining some parts of the text of this article, though all information presented remains the result of the authors' own research and contributions. The authors acknowledge support by the Helmholtz Foundation program "Materials System Engineering" (MSE) at the Karlsruhe Institute of Technology. M.T. acknowledges support by the DACStorE project, funded by the Initiative and Networking Fund of the Helmholtz Association (grant agreement number KA2-HSC-12).

Open access funding enabled and organized by Projekt DEAL.

## Conflict of Interest

The authors declare no conflict of interest.

## Data Availability Statement

The data that support the findings of this study are available through a repository.

## Keywords

layer-by-layer, machine learning, roughness-optimization, SURMOF

Received: March 17, 2024

Revised: May 16, 2024

Published online:

- [1] O. M. Yaghi, G. Li, H. Li, *Nature* **1995**, 378, 703.
- [2] H.-C. Zhou, S. Kitagawa, *Chem. Soc. Rev.* **2014**, 43, 5415.
- [3] S. Kitagawa, R. Kitaura, S.-i. Noro, *Angew. Chem., Int. Ed.* **2004**, 43, 2334.
- [4] J. L. C. Rowsell, O. M. Yaghi, *Microporous Mesoporous Mater.* **2004**, 73, 3.
- [5] J. Liu, B. Lukose, O. Shekhah, H. K. Arslan, P. Weidler, H. Gliemann, S. Bräse, S. Grosjean, A. Godt, X. Feng, K. Müllen, I.-B. Magdau, T. Heine, C. Wöll, *Sci. Rep.* **2012**, 2, 921.
- [6] M. Kalaj, S. M. Cohen, *ACS Cent. Sci.* **2020**, 6, 1046.
- [7] Z. Wang, J. Liu, H. K. Arslan, S. Grosjean, T. Hagendorn, H. Gliemann, S. Bräse, C. Wöll, *Langmuir* **2013**, 29, 15958.
- [8] S. Begum, Z. Hassan, S. Bräse, C. Wöll, M. Tsotsalas, *Acc. Chem. Res.* **2019**, 52, 1598.
- [9] R. Freund, O. Zaremba, G. Arnauts, R. Ameloot, G. Skorupskii, M. Dincă, A. Bavykina, J. Gascon, A. Ejsmont, J. Goscianska, M. Kalmutzki, U. Lächelt, E. Ploetz, C. S. Diercks, S. Wuttke, *Angew. Chem., Int. Ed.* **2021**, 60, 23975.
- [10] R. Haldar, L. Heinke, C. Wöll, *Adv. Mater.* **2020**, 32, 1905227.
- [11] O. Shekhah, H. Wang, S. Kowarik, F. Schreiber, M. Paulus, M. Tolan, C. Sternemann, F. Evers, D. Zacher, R. A. Fischer, C. Wöll, *J. Am. Chem. Soc.* **2007**, 129, 15118.
- [12] D.-H. Chen, H. Gliemann, C. Wöll, *Chem. Phys. Rev.* **2023**, 4, 011305.
- [13] S. Wuttke, D. D. Medina, J. M. Rotter, S. Begum, T. Stassin, R. Ameloot, M. Oschatz, M. Tsotsalas, *Adv. Funct. Mater.* **2018**, 28, 1801545.
- [14] B. Hosseini Monjezi, K. Kutonova, M. Tsotsalas, S. Henke, A. Knebel, *Angew. Chem., Int. Ed.* **2021**, 60, 15153.
- [15] S. Okur, T. Hashem, E. Bogdanova, P. Hodapp, L. Heinke, S. Bräse, C. Wöll, *ACS Sens.* **2024**, 9, 622.
- [16] L. Heinke, Z. Gu, C. Wöll, *Nat. Commun.* **2014**, 5, 4562.
- [17] B. Mohan, S. Kumar, H. Xi, S. Ma, Z. Tao, T. Xing, H. You, Y. Zhang, P. Ren, *Biosens. Bioelectron.* **2022**, 197, 113738.
- [18] S.-I. Ohira, Y. Miki, T. Matsuzaki, N. Nakamura, Y.-k. Sato, Y. Hirose, K. Toda, *Anal. Chim. Acta* **2015**, 886, 188.
- [19] I. Stassen, N. Burtch, A. Talin, P. Falcaro, M. Allendorf, R. Ameloot, *Chem. Soc. Rev.* **2017**, 46, 3185.
- [20] L. S. Xie, G. Skorupskii, M. Dincă, *Chem. Rev.* **2020**, 120, 8536.
- [21] H. Yuan, N. Li, W. Fan, H. Cai, D. Zhao, *Adv. Sci.* **2022**, 9, 2104374.
- [22] M.-S. Yao, W.-H. Li, G. Xu, *Coord. Chem. Rev.* **2021**, 426, 213479.
- [23] P. Li, Z. Zhou, Y. S. Zhao, Y. Yan, *Chem. Commun.* **2021**, 57, 13678.
- [24] A. Zuliani, N. Khair, C. Carrillo-Carrión, *Anal. Bioanal. Chem.* **2023**, 415, 2005.
- [25] J. F. Olorunyomi, S. T. Geh, R. A. Caruso, C. M. Doherty, *Mater. Horiz.* **2021**, 8, 2387.
- [26] Z.-G. Gu, A. Pfriend, S. Hamsch, H. Breitwieser, J. Wohlgemuth, L. Heinke, H. Gliemann, C. Wöll, *Microporous Mesoporous Mater.* **2015**, 271, 82.
- [27] M. Oldenburg, A. Turshatov, D. Busko, S. Wollgarten, M. Adams, N. Baroni, A. Welle, E. Redel, C. Wöll, B. S. Richards, I. A. Howard, *Adv. Mater.* **2016**, 28, 8477.
- [28] K. Möller, K. Fink, L. Schöttner, M. Koenig, L. Heinke, C. Wöll, *ACS Appl. Mater. Interfaces* **2017**, 9, 37463.
- [29] W. Wang, D. I. Sharapa, A. Chandresh, A. Nefedov, S. Heißler, L. Heinke, F. Studt, Y. Wang, C. Wöll, *Angew. Chem., Int. Ed.* **2020**, 59, 10514.
- [30] P. Kalhor, N. Jung, S. Bräse, C. Wöll, M. Tsotsalas, P. Friederich, *Adv. Funct. Mater.* **2023**, 34, 2302630.
- [31] L. Pilz, C. Natzeck, J. Wohlgemuth, N. Scheuermann, S. Spiegel, S. Oßwald, A. Knebel, S. Bräse, C. Wöll, M. Tsotsalas, N. Prasetya, *J. Mater. Chem. A* **2023**, 11, 24724.
- [32] Y. Luo, S. Bag, O. Zaremba, A. Cierpka, J. Andreo, S. Wuttke, P. Friederich, M. Tsotsalas, *Angew. Chem., Int. Ed.* **2022**, 61, 202200242.
- [33] S. S.-Y. Chui, S. M.-F. Lo, J. P. H. Charmant, A. G. Orpen, I. D. Williams, *Science* **1999**, 283, 1148.
- [34] L. Pilz, C. Natzeck, J. Wohlgemuth, N. Scheuermann, P. G. Weidler, I. Wagner, C. Wöll, M. Tsotsalas, *Adv. Mater. Interfaces* **2023**, 10, 2201771.
- [35] S. M. Moosavi, A. Chidambaram, L. Talirz, M. Haranczyk, K. C. Stylianou, B. Smit, *Nat. Commun.* **2019**, 10, 539.
- [36] J. Liu, C. Wöll, *Chem. Soc. Rev.* **2017**, 46, 5730.
- [37] C.-L. Lin, P.-C. Huang, S. Graessle, C. Grathwol, P. Tremouilhac, S. Vanderheiden, P. Hodapp, S. Herres-Pawlis, A. Hoffmann, F. Fink, G. Manolikakes, T. Opatz, A. Link, M. M. B. Marques, L. J. Daumann, M.

Tsotsalas, F. Biedermann, H. Mutlu, E. Täuscher, F. Bach, T. Drees, S. Neumann, N. Jung, S. Bräse, Supporting Sustainability of Chemistry by Linking Research Data with Physically Preserved Research Materials, preprint, Chemistry, **2023**, <https://chemrxiv.org/engage/chemrxiv/article-details/64a2fa0cba3e99daef73fa6a>.

- [38] K. Möller, J. Singh Malhi, J. Wohlgemuth, R. A. Fischer, C. Wöll, H. Gliemann, L. Heinke, *Dalton Trans.* **2018**, 47, 16474.
- [39] M. Kjærviik, P. M. Dietrich, A. Thissen, J. Radnik, A. Nefedov, C. Natzeck, C. Wöll, W. E. S. Unger, *J. Electron Spectrosc. Relat. Phenom.* **2021**, 247, 147042.