

Electrosynthetic Screening and Modern Optimization Strategies for Electrosynthesis of Highly Value-added Products

Maurice Dörr⁺,^[a] Maximilian M. Hielscher⁺,^[a] Jonny Proppe,^[b] and Siegfried R. Waldvogel^{*[a]}

In memory of Christoph Naethbohm

Unlike common analytical techniques such as cyclic voltammetry, statistics-based optimization tools are not yet often in the toolbox of preparative organic electrochemists. In general, experimental effort is not optimally utilized because the selection of experimental conditions is based on the one-

variable-at-a-time principle. We will summarize statistically motivated optimization approaches already used in the context of electroorganic synthesis. We discuss the central ideas of these optimization methods which originate from other fields of chemistry in relation to electrosynthetic applications.

1. Introduction

Organic electrosynthesis enjoys a revival, growing interest, and enthusiasm within the organic chemical community.^[1–4] When the concepts of *green chemistry* are carefully applied, electrochemistry enables highly sustainable and eco-friendly synthetic strategies.^[5] The electrochemical activation is of particular interest to electrify organic synthesis^[6] and producing value-added products.^[7] Consequently, electroorganic synthesis is considered as 21st century technique and can represent a cost-efficient method.^[8] Besides the green aspect the electroorganic synthesis can substantially short cut the number of synthetic steps being required.^[9,10] Great efforts are made to bring this field into the toolkit of every chemist. The expansion of the scope of possible applications for electrosynthesis progresses rapidly.^[2–4,11] We are striving to streamline the experimental procedure itself, enabling others to easily adopt it, since many parameters are often not reported to provide good reproducibility.^[2,12] But still the optimization remains challenging in many cases and attention must be paid to many variables of the studied system. More precisely, the already big parameter space of conventional organic synthesis is further enlarged by addition of electrolysis specific variables like the amount of applied charge and the current density. Besides that, the impact of sometimes considered minor factors like the stirring rate or

inter electrode distance might be surprisingly big in some cases due to the electrolysis' nature as an interfacial reaction.^[12–14]

The resulting complexity creates the desire for systematic optimization strategies that are capable of handling these mostly non-linear systems we are confronting. When a *One Variable at a Time* approach is insufficient, a common step forward is using *Design of Experiments* (DoE) techniques that include building a regression model, thus giving detailed information about the system under investigation.^[14–19] If multiple combinations of substrates are to be considered, including measured and calculated molecular parameters can be very achieving but requires the use of more sophisticated approaches like predictive modeling methods (e.g. *Multivariate Linear Regression* or *Principal Component Analysis*) and machine learning.^[20–22]

The mentioned methods require balanced experimental data to produce good performing models. So, chemists must tackle the paradigm of neglecting negative results and make the same effort for bad as for good reactions to obtain high quality data for both cases.^[23]

Automated setups push the boundaries of quick generation of data and flow-operation is preferred over batch mode in this case. This is due to the relatively simple inclusion of inline measurements, which allow for a real-time feedback for the outcome of the reaction.^[24] Using automated flow setups in classical organic synthesis led to the development of highly autonomous synthetic tools from self-optimizing setups using active learning or optimization algorithms to fully AI driven synthesis robots.^[25–28] In this article we want to sum up how these elaborate strategies can be applied to electrochemistry for accelerated development of electrosynthetic pathways towards highly value-added products.

2. Electrochemical Parameters

The parameter space for electroorganic reactions includes not only parameters for classical organic synthesis, but also those

[a] M. Dörr,⁺ M. M. Hielscher,⁺ Prof. Dr. S. R. Waldvogel
Department of Chemistry
Johannes Gutenberg University
Duesbergweg 10–14, 55128 Mainz, Germany
E-mail: waldvogel@uni-mainz.de

[b] Dr. J. Proppe
Institute of Physical Chemistry
Georg-August Universität
Tammannstr. 6, 37077 Göttingen, Germany

[⁺] These authors contributed equally and are listed alphabetically.

 © 2021 The Authors. ChemElectroChem published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution Non-Commercial NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

required to describe the electrodes and the electrolyte system. Including continuous and directly controllable parameters like concentrations, ratios, temperature, or pressure as well as categorical (non-continuous) parameters like the choice of solvent, nature of the electrode material, reactants, or the reactor setup.

Every mode of operation brings its own set of parameters: In general, one must consider the electrode materials, their geometry, their size, the distance between working- and counter-electrode and, if used, the supporting electrolyte and its concentration. In galvanostatic mode, the current density and amount of charge are added to the list and in potentiostatic mode the voltage needs to be considered. An overview of the parameters is given in Figure 1. Note that in some scenarios (e.g. virtual screening or predictive modeling^[20,21]) it is very rewarding to consider additional supportive descriptors. The latter might include information about electrolyte, electrode material or intermediates involved in the reaction.

3. Optimization Strategies

In a fair number of cases optimization procedures follow the *One Variable at a Time* (OVAT) principle. The experimenter selects a start set of reaction conditions and optimizes step by step, parameter for parameter. The idea behind this is already anchored in school physics, just change one setting to know where the systems change comes from, if there is any. This brings up two problems, first what set of settings does one use

to start and less obviously, how does one quantify interactions between parameters. The former can be dealt with by prior knowledge, whereas the latter can only be remedied by a more sophisticated optimization procedure.

In many cases, efficient optimization procedures can lead to a better understanding of the investigated system with the least number of experiments needed. Although the presented tools do not give a direct insight into the mechanistic course, they provide hints which parameters play significant or even decisive roles. This knowledge can be used to sharpen mechanistic proposals and deepen the understanding.

3.1. Design of Experiments

Design of Experiments (DoE) refers to a series of (experimental) design patterns that help fulfilling these requirements and aim providing an empirical model for the studied system. The patterns are designed in such a way that for a given number of parameters of a process to be investigated, the respective cause-effect relationships with respect to one or more process responses are determined with minimal experimental effort by efficiently arranging the experimental points in the parameter space. Their application encourages the experimenter to consider various methods to maximize the amount of information obtained while minimizing the influence of sources of error.^[29,30] It may seem difficult at first to apply DoE to one's own optimization problem but it turns out very rewarding afterwards and there are many helpful and good introductions



Maurice Dörr received his B.Sc. degree in chemistry from Johannes Gutenberg-University Mainz working on anodic C–C cross-coupling reactions in 2016 and his M.Sc. working on anodic C–N cross-coupling reactions in 2018 supervised by Prof. Dr. Siegfried R. Waldvogel. Currently, he is engaged in the application of Design of Experiments (DoE) towards electro-organic synthesis as a graduate student in the Waldvogel lab.



Maximilian Hielscher obtained his B.Sc. degree in chemistry from Johannes Gutenberg-University Mainz for studies on Mo(V)-mediated oxidative domino coupling and his M. Sc. degree for evaluating statistical optimization methods in electrochemical oxidative phenol-arene C–C cross-coupling under the supervision of Prof. Dr. Siegfried R. Waldvogel. He is currently working on his PhD research project on the optimization of anodic cross-coupling reactions in flow electrolyzers as part of the DFG research project UNODE.



Jonny Proppe studied chemistry in Hamburg and obtained his doctorate (Dr. sc.) in 2018 from ETH Zürich (Prof. M. Reiher), for which he was awarded the IBM Research prize in 2020. He was an Early Postdoc.Mobility fellow of the Swiss National Science Foundation at Harvard University and the University of Toronto (Prof. A. Aspuru-Guzik). Currently, he is a research associate of the DFG-funded RTG 2455 "BENCH" at Georg-August University, Göttingen. His research interests revolve around the development of data-driven methods to accelerate chemical innovation and knowledge production.



Siegfried R. Waldvogel studied chemistry in Konstanz and received his Ph.D. in 1996 from University of Bochum/Max Planck Institute for Coal Research (Prof. M. T. Reetz). After post-doctoral research in La Jolla, CA (Prof. J. Rebek, Jr.), he started his own research at the Universities of Münster and Bonn. He became full professor for organic chemistry at Johannes Gutenberg University Mainz in 2010. His research interests are novel electro-organic transformations including bio-based feedstocks from electrosynthetic screening to scale-up. In 2018, he cofounded ESy-Labs GmbH, which provides custom electrosynthesis and contract R&D.

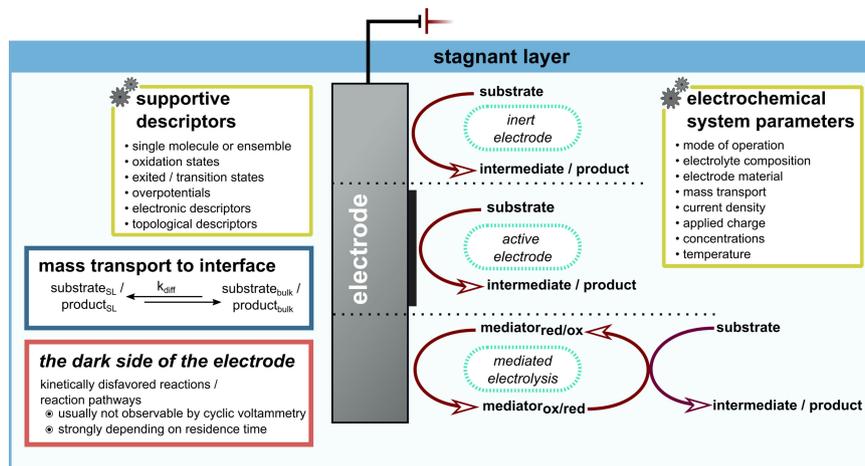


Figure 1. Common parameters and pathways in electrosynthetic processes. The system parameters are directly controllable by the experimenter and should be considered during the optimization process. Supportive descriptors are useful to consider when large data sets are sought for (e.g. in machine learning), they may also include parameters that describe the electrode material or components of the electrolyte.

in literature.^[29–31] While there are commercial software packages on the market, there are also free Python or R projects that have implemented the common designs and evaluations (e.g. *pyDOE*^[32] the *DoE R-Plugin*^[33]).

Factorial designs are commonly used and easy to establish. In their simplest form they comprise two levels for each parameter and include every possible combination of the settings of all parameters (two level full factorial design). To tackle the large number of experiments in factorial designs with big parameter sets, fractions of the full factorial design (fractional factorial design) can often be used with minor impact on the information obtained from the resulting model. In addition to factorial designs there are designs that allow for the application of models including higher order polynomial terms. Central composite designs for example, may be constructed from an initial factorial design to provide that feature.

While DoE has been successfully used as an optimization tool in many areas of chemistry and has found a firm place in the industry in the Six Sigma environment, it is underrepresented in electroorganic synthesis. A representative selection of DoE applications in electroorganic synthesis is shown in Figure 2. *Hilt* and co-workers published their first DoE-based optimization in 2018, in which they studied and optimized anodic iodination reactions.^[15] Although the initial linear optimization quickly produced outstanding yields for the example substrate, only moderate yields could be achieved for their aimed substrate scope. By using a central composite experimental design, they investigated three reaction parameters (solvent mixture ratio, KI stoichiometry and applied charge) in a total of 18 experiments. The data evaluation revealed a second order interaction between the applied amount of charge and the concentration of methanol, which is due to overoxidation of the solvent at high methanol concentration and excess charge. The optimized conditions allow the electrochemical conversion of various electron-rich and electron-poor substituted arenes in good to very good yields.

In 2019 *Wirth* and co-workers combined their electrochemical flow cell with online 2D HPLC analysis in a DoE driven optimization approach.^[17] They exploited the memory effect of chirality in *N*-aryl carbonylated amino acid derivatives under non-Kolbe electrolysis conditions to generate enantiomerically enriched alkoxyated amides via an acyliminium ion intermediate. The usage of desirability functions allowed a simultaneous optimization of both targets, based on the same experimental data set. They started with a fast linear screening of electrode materials, substrate concentrations and flow rates, followed by a fractional 2^{5-1} design to cover the influence of the concentration of starting material, applied charge, temperature, flow rate, and—as a categorical parameter—the anode material (graphite and glassy carbon). In a second step they used a full factorial 2^3 design to fine tune flow rate, amount of charge and temperature ending up with optimized reaction conditions with yields up to 100% and enantioselectivities up to 70% *ee*.

Our group lately used a DoE driven optimization approach to investigate the anodic C–C-cross coupling of 2-methoxy-4-methylphenol and 2,4-dimethylphenol in flow electrolyzers.^[19] This transformation has been studied by our group for quite some time in an OVAT approach, successfully scaling up to larger electrolyzers.^[34] However, the optimization of the yield has always been difficult due to the high number of reaction parameters involved. We then investigated the reaction in a fractional factorial 2^{7-3} experimental design, following the results of linear optimization with respect to the parameter limits. Data analysis revealed the presence of several two-factor interactions that had complicated the OVAT optimization. The main effect plots indicated an independence of current density and starting concentration in the studied region, which allowed us to almost triple the space-time yield, with an overall yield of up to 85%.

For the optimization of an anodic, salt-free synthesis of aryl HFIP ethers our group again chose a DoE based approach. The results were compared to the conditions optimized in an OVAT fashion. For many of the substrates under investigation the new DoE based conditions led to higher yields up to 59%, while

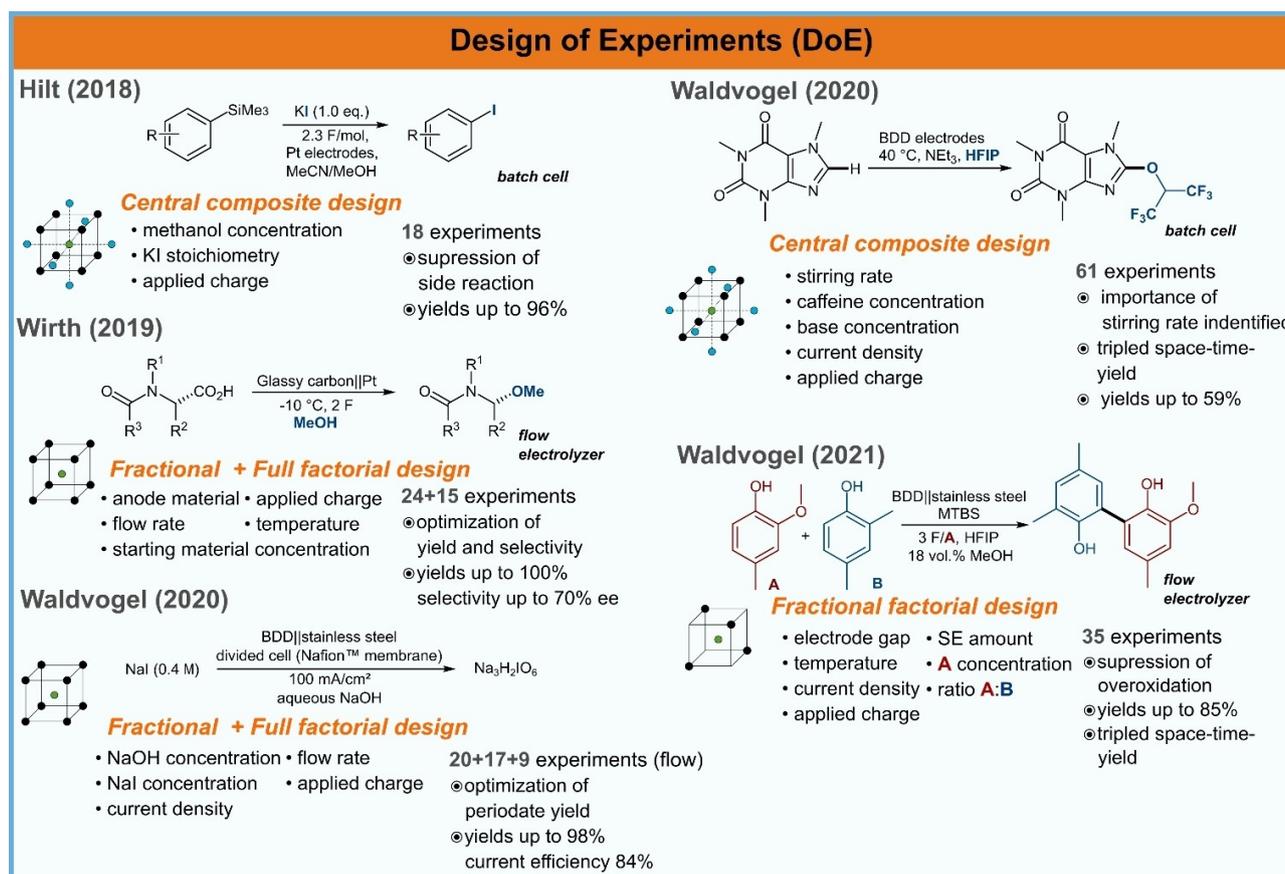


Figure 2. Applications of Design of Experiments in electrosynthesis.^[14,15,17,19]

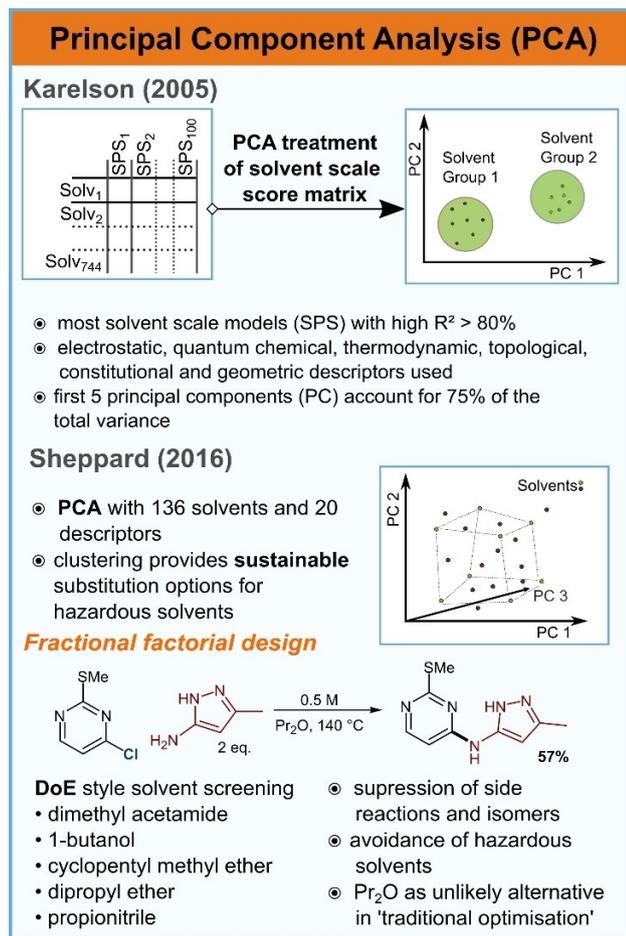
dropping the reaction times to about a third giving a more convenient electrolysis overall. The application of factorial designs in this case also revealed the stirring rate as a crucial factor for the reactions outcome.^[14,35]

Furthermore, our group used a DoE supported optimization for the electrochemical oxidation of sodium iodide to periodate at a BDD anode.^[36] Previous methods of producing periodate are based on the use of lead dioxide anodes, but they are known to slowly decompose into solution during oxidation. Therefore, expensive purification of periodate is necessary before use as reagent for sensitive applications. Consequently, a carbon anode-based alternative was highly desirable. In the optimization of batch experiments, higher current density was shown to have a positive effect on yield, but a full-factorial 2⁴ design for scale up only achieved a yield of 60%. A corresponding optimization of the reaction in a 12 cm² divided flow cell could be transferred to a cell with 48 cm² anode area without affecting the yield. Up to 98% *para*-periodate was obtained at a current density of 100 mA/cm². The obtained periodate was suitable for the shortest syntheses of the anti-epileptic drug levetiracetam.^[10]

3.2. Principal Component Analysis

When working with complex systems and trying to consider every possibly influential variable one might be confronted with massive datasets. In this case a principal component analysis might be the right tool for the job. In very simplified terms, principal component analysis is an approach to reduce the dimensionality of a given dataset by construction of a set of principal components (PC) as linear combinations of the variables contained in the original dataset. In this process, the linear combinations are built to be as meaningful as possible, i.e. they describe the variance as sufficient as possible. Simple to use implementations of principle component analysis are included in *scikit-learn* and *statsmodels* for example.^[37,38]

An illustrative example is given from the field of quantitative structure-property relationship (QSPR). The group of Karelson and co-workers investigated a list of 774 solvents and 127 solvent scales (Figure 3).^[39] The solvent scales used here are multilinear regression equations as a function of descriptors describing the solvents under consideration. The descriptors come from different classes, such as constitutional, thermodynamic or quantum chemical descriptors. They used a combinatorial approach, utilizing existing experimental data and calculated descriptor values. For each solvent molecule they determined the score on each solvent scale, ending up with a



descriptors are determined, for example, in quantum mechanical calculations, and the performance of the ligand is predicted using the regression model.^[52] In the laboratory, corresponding complexes are then synthesized from all performance ranges and the model is evaluated on the basis of the results. It is of great importance not only to test “good” catalysts or “good reaction conditions”, but to test the model in its entire prediction range.^[20] Correlation plots between quantum-chemically obtained descriptors and measured properties can give an indication of different or similar system behavior. For example, *Nematollahi* and co-workers were able to show an identical electron transfer mechanism by linear relation of the HOMO energies with the measured half wave oxidation potentials of different catechol derivatives.^[53]

If this concept is transferred to organic electrochemistry, properties of the substrate in the form of physical or calculated descriptors could be used for virtual screening.

3.3.2. Multivariate Linear Regression

The group of *Hilt* and co-workers used multivariate linear regression in the optimization of the electrochemical trifluoroethoxylactonisation (Figure 4).^[54] The aim of the study was to obtain an experimentally inexpensive approach to possible chiral mediators by multivariate modeling. Initially, they investigated the parameter space of a test reaction using iodobenzene as a mediator in a D-optimal screening design in 25 runs. This allowed them to increase the yield of the test reaction from an initial 27% to 78%. In the next step, they tested the performance of different commercially available mediators using the optimized conditions of the test reaction and built a regression model using the yields and theoretically calculated descriptors of the mediators. They calculated descriptor values for several more difficult to synthesize chiral mediators, whose performance they predicted based on the MLR model. Validation showed good predictive power of the model and the first chiral mediator test showed a promising yield of 56%. This doubled the yield previously known for this chiral mediator.

3.3.3. Nonlinear Methods

The group of *Ding* and co-workers recently used cyclic voltammetry (CV) to generate descriptors for electro-organic synthesis (Figure 5). They used the onset potential, Tafel slope and effective oxidation potentials to build an “electro-descriptor diagram”. While CV studies are reasonably quick to conduct, it still takes considerable experimental effort, thus forcing *Ding* et al. to work on relatively small datasets up to 29 entries including the test set. Machine Learning is especially appealing when descriptors can be quickly and inexpensively generated. Nevertheless, they applied a decision tree and a *k*-nearest neighbors algorithm to the obtained data from the model reactions, using 60% of the data as a training set and 40% for testing. The chosen methods are typically used for classification

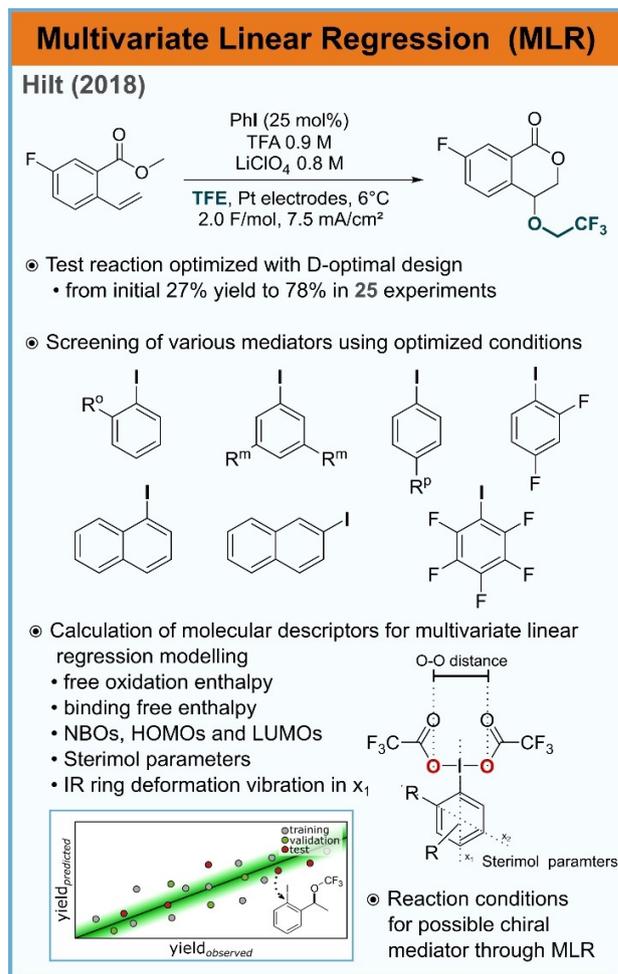


Figure 4. Workflow of *Hilt* for virtual screening of various iodine mediators for the electrochemical trifluoroethoxylactonization reaction. First, the test reaction was optimized using iodophenol by means of DoE. With the optimized conditions, different iodine mediators were tested and linked to quantum chemical descriptors by the means of MLR. A possible chiral mediator was used in the reaction according to its descriptor-dependent scoring, doubling the literature-known yield for the latter.^[54]

tasks and incorporate different strategies to solve them.^[40] They demonstrated that the electro-descriptors could be used to identify a hot zone leading to good reaction outcomes and they were able to separate reactive and nonreactive species.^[55] It should be noted that except for the case of TEMPO-mediated oxidative dehydrogenation of N-heterocycles, for preparative scale electrolysis a constant potential was used. In the economically more interesting form, electrolysis at constant current, the potential is raised to the desired current flow. This complicates the modeling, since the oxidation and reduction potentials of the other substrates and the reaction products, as well as the kinetics of these electron transfer reactions and the subsequent reactions must also be considered. Furthermore, CV only maps the situation at the electrode, but is largely oblivious to subsequent reactions in the bulk or at the transition to the stagnant layer.^[56]

Modestino and co-workers successfully used an artificial neural network to optimize a pulsed electrolysis of acrylonitrile

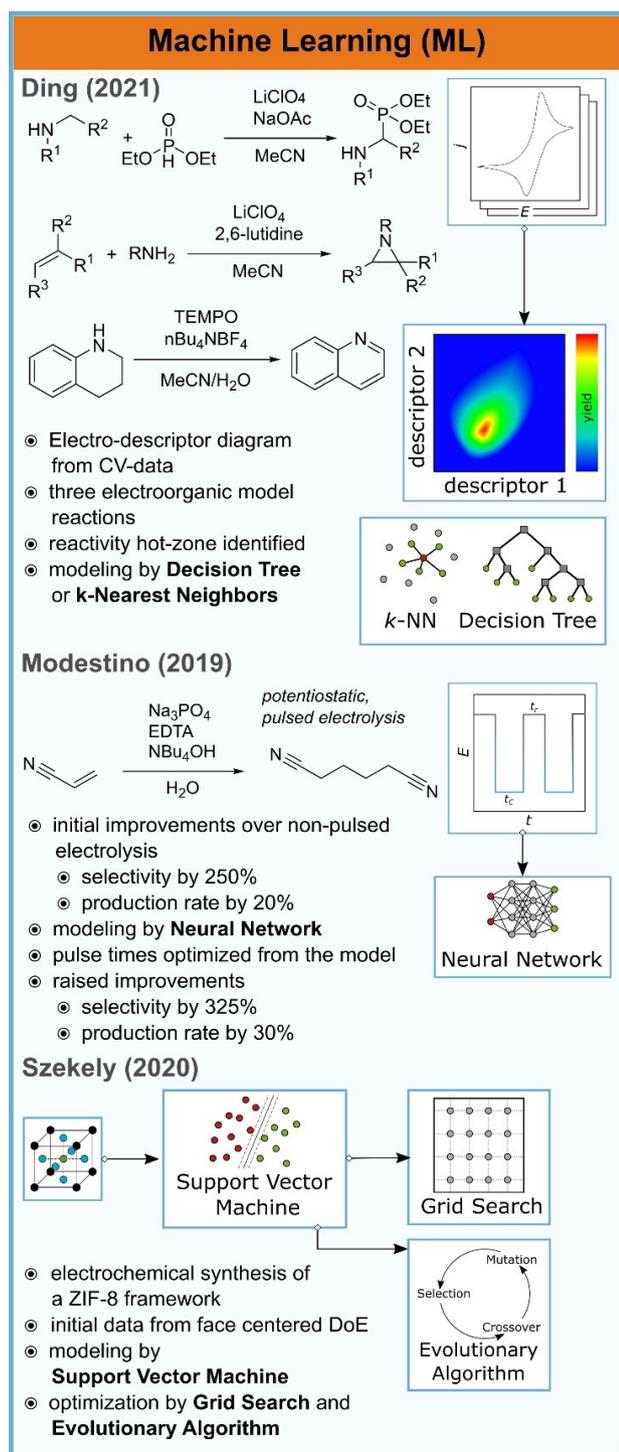


Figure 5. Examples for the application of machine learning in electrochemistry.^[55,57,58]

regarding the production rate of adiponitrile and the selectivity of the reaction (Figure 5). This class of algorithms is certainly the most popular today. This is due to its great versatility, flexibility and state of the art performance in finding complex patterns. Though, it should be noted that the application of a neural network requires careful handling of several hyperparameters and that they tend to overfit easily if not used

properly.^[40] The initial optimization in an OVAT fashion led to an increase in selectivity of 250% and production rate of 20% compared to the non-pulsed electrolysis, proving the superiority of their process. Applying a 2-layer feed-forward neural network with 9 hidden neurons, the improvements could be raised even further to 325% and 30%, respectively. Thus, impressively illustrating the networks use in the optimization process even though only a small dataset of 16 data points was used. It should be noted again that the received model has to be critically inspected when working with these small data sets to ensure reasonable generalization.^[40] The data was split to 70% training set, 15% test set and 15% validation set.^[57]

When using elaborate optimization strategies like the ones described here, it is simple to extend the targeted goals by defining a desirability function that contains the weighted measures sought to be optimized. Szekely and co-workers used this approach to design a sustainable electrochemical process for the synthesis of a ZIF-8 metal-organic framework (Figure 5). They included measures for product quality and yield as well as the process' sustainability (carbon footprint, E-factor, energy consumption) in the desirability function. A Support Vector Machine (SVM) was used for modelling with the experimental data of an initially run DoE. The data set obtained via the DoE consisted of 27 runs and cross validation was used to tune the SVMs hyperparameters. The resulting model was then used for parameter optimization. Two approaches were compared, one running grid search on the model, the other using an evolutionary algorithm. The constructed desirability function was used as the figure of merit.^[58]

3.3.4. Automation and Autonomous Optimization

In other areas of synthetic chemistry, machine learning techniques have been used to greater extent.^[22,28,42] Jensen and co-workers accomplished almost full automation of a robotic flow synthesis platform for organic compounds. The whole process included several artificial neural networks and a Monte Carlo tree search algorithm for synthesis planning.^[25] Using literature data for modelling in this big data fashion requires one to carefully consider the bias in literature toward reporting positive results only.^[25,59] The reactor used by Jensen et al. is modular and reconfigured by a robotic arm, leaving only some refining during the recipe formulation to the user.^[25] It should be highlighted that, if applicable, flow operation is the preferred mode for a high degree of automation and online reaction monitoring in this case.^[24,26–28] The latter was not implemented by Jensen et al. due to their work's focus on synthesis planning and not the optimization of a given reaction.^[25]

Implementation of a feedback loop enables autonomous optimization procedures using active learning algorithms. Felipin and co-workers used a simplex algorithm to autonomously optimize the C–H arylation of indole-3-acetic acid derivatives reaching yields up to 89%.^[28] The Doyle group recently demonstrated the use of Bayesian optimization algorithms for chemical reactions in a striking way. After initial development

and tuning on experimental data from literature on two palladium-catalyzed cross-coupling reactions, they collected results from 1728 HTE experiments covering the complete search space previously defined. Again, a palladium-catalyzed cross-coupling reaction was used for this part of their studies. The whole experimental dataset was then used to develop a quiz game to collect input by chemists with it. Afterwards the human optimization attempts and the Bayesian optimizer were compared, showing that both performance and variance of the algorithm were superior. While no truly autonomous optimization was employed here, it was simulated on the already known results from the HTE campaign.^[60]

4. Conclusion

We see great potential in the application of modern optimization methods in organic electrosynthesis. Simple OVAT approaches just seem inappropriate for the complex nature of electroorganic conversions. DoE helps to understand cause-effect relationships and multiparameter interactions not resolved in OVAT. It also offers adequate solutions for systems with slightly higher amounts of parameters. When working with large datasets principal component analysis is a handy tool to narrow down the dimensionality of the parameter space while still maintaining reasonably accurate representations of the higher dimensional space. Thus, PCA can help to systematically include solvents in optimization methods as described above. Predictive modeling can be used to broaden the optimization scope by incorporating substrate derived descriptor data. Using machine learning opens countless other options from computer assisted synthesis planning to fully autonomous self-optimizing flow reactors.

In many cases, combinations of the described methods can be found in literature since there are manifold challenges in chemistry that sometimes ask for multiple steps and different tools to appropriately solve them. In order to apply these approaches, a paradigm shift has to happen. The information of both, well working as well as inferior experiments is needed for balanced data sets and therefore well performing models. Therefore, the results for both have to be of the same quality. Well, efficient working synthetic chemists often discard the runs if they see it is not well working. Therefore, old sets of experiments cannot easily be exploited for these statistical methods, because half of the information is already lost. However, with these modern statistical tools the field of electroorganic synthesis will be propelled into a new era of time and cost-efficient screening and optimization.

Acknowledgements

The authors thank the Carl Zeiss Foundation for funding to the ECHELON project. Support by the State of Rhineland-Palatinate with the center SusInnoScience is highly appreciated. M. Hielscher thanks the DFG for founding in the frame of the research training group GRK2516. Tom Wirtanen is thanked for

many fruitful discussions. J.P. acknowledges funding by the German Research Foundation (DFG) via project 389479699/GRK2455. Open access funding enabled and organized by Projekt DEAL.

Conflict of Interest

The authors declare no conflict of interest.

Keywords: electrochemistry · optimization · statistical methods · machine learning · chemoinformatics

- [1] S. R. Waldvogel, B. Janza, *Angew. Chem. Int. Ed.* **2014**, *53*, 7122–7123; *Angew. Chem.* **2014**, *126*, 7248–7249.
- [2] M. Yan, Y. Kawamata, P. S. Baran, *Chem. Rev.* **2017**, *117*, 13230–13319.
- [3] J. L. Röckl, D. Pollok, R. Franke, S. R. Waldvogel, *Acc. Chem. Res.* **2020**, *53*, 45–61.
- [4] M. D. Kärkäs, *Chem. Soc. Rev.* **2018**, *47*, 5786–5865.
- [5] a) Y. Yuan, A. Lei, *Nat. Commun.* **2020**, *11*, 802; b) P. Anastas, N. Eghbali, *Chem. Soc. Rev.* **2010**, *39*, 301–312.
- [6] A. Wiebe, T. Gieshoff, S. Möhle, E. Rodrigo, M. Zirbes, S. R. Waldvogel, *Angew. Chem. Int. Ed.* **2018**, *57*, 5594–5619; *Angew. Chem.* **2018**, *130*, 5694–5721.
- [7] S. Möhle, M. Zirbes, E. Rodrigo, T. Gieshoff, A. Wiebe, S. R. Waldvogel, *Angew. Chem. Int. Ed.* **2018**, *57*, 6018–6041; *Angew. Chem.* **2018**, *130*, 6124–6149.
- [8] a) D. Pollok, S. R. Waldvogel, *Chem. Sci.* **2020**, *11*, 12375–12592; b) J. Seidler, J. Strugatchi, T. Gärtner, S. R. Waldvogel, *MRS Energy Sustainability* **2020**, *7*.
- [9] a) S. P. Blum, T. Karakaya, D. Schollmeyer, A. Klapars, S. R. Waldvogel, *Angew. Chem. Int. Ed.* **2021**, *60*, 5056–5062; *Angew. Chem.* **2021**, *133*, 5114–5120; b) E. J. Horn, B. R. Rosen, P. S. Baran, *ACS Cent. Sci.* **2016**, *2*, 302–308; d) K. D. Moeller, *Tetrahedron* **2000**, *56*, 9527–9554.
- [10] S. Arndt, B. Grill, H. Schwab, G. Steinkellner, U. Pogorevcnik, D. Weis, A. M. Nauth, K. Gruber, T. Opatz, K. Donsbach et al., *Green Chem.* **2021**, *23*, 388–395.
- [11] a) R. D. Little, K. D. Moeller, *Chem. Rev.* **2018**, *118*, 4483–4484; b) S. R. Waldvogel, S. Lips, M. Selt, B. Riehl, C. J. Kampf, *Chem. Rev.* **2018**, *118*, 6706–6765; c) E. J. Horn, B. R. Rosen, P. S. Baran, *ACS Cent. Sci.* **2016**, *2*, 302–308; d) K. D. Moeller, *Tetrahedron* **2000**, *56*, 9527–9554.
- [12] S. Beil, D. Pollok, S. R. Waldvogel, *Angew. Chem. Int. Ed.* **2021**, in press, doi.org/10.1002/anie.202014544; *Angew. Chem.* **2021**, in press, doi.org/10.1002/ange.202014544.
- [13] a) Y. Mo, Z. Lu, G. Rughoobur, P. Patil, N. Gershenfeld, A. I. Akinwande, S. L. Buchwald, K. F. Jensen, *Science* **2020**, *368*, 1352–1357; b) T. Gieshoff, V. Trieu, J. Heijl, S. R. Waldvogel, *Beilstein J. Org. Chem.* **2018**, *14*, 1578–1582.
- [14] M. Dörr, J. L. Röckl, J. Rein, D. Schollmeyer, S. R. Waldvogel, *Chem. Eur. J.* **2020**, *26*, 10195–10198.
- [15] R. Möckel, J. Hille, E. Winterling, S. Weidemüller, T. M. Faber, G. Hilt, *Angew. Chem. Int. Ed.* **2018**, *57*, 442–445; *Angew. Chem.* **2018**, *130*, 450–454.
- [16] P. M. Murray, F. Bellany, L. Benhamou, D.-K. Bučar, A. B. Tabor, T. D. Sheppard, *Org. Biomol. Chem.* **2016**, *14*, 2373–2384.
- [17] M. Santi, J. Seitz, R. Cicala, T. Hardwick, N. Ahmed, T. Wirth, *Chem. Eur. J.* **2019**, *25*, 16230–16235.
- [18] a) S. A. Weissman, N. G. Anderson, *Org. Process Res. Dev.* **2015**, *19*, 1605–1633; b) G. Hanrahan, K. Lu, *Crit. Rev. Anal. Chem.* **2006**, *36*, 141–151.
- [19] M. M. Hielscher, B. Gleede, S. R. Waldvogel, *Electrochim. Acta* **2021**, *368*, 137420.
- [20] A. G. Maldonado, G. Rothenberg, *Chem. Soc. Rev.* **2010**, *39*, 1891–1902.
- [21] J. L. Melville, E. K. Burke, J. D. Hirst, *Comb. Chem. High Throughput Screening* **2009**, *12*, 332–343.
- [22] A. Tkatchenko, *Nat. Commun.* **2020**, *11*, 4125.
- [23] P. Raccuglia, K. C. Elbert, P. D. F. Adler, C. Falk, M. B. Wenny, A. Mollo, M. Zeller, S. A. Friedler, J. Schrier, A. J. Norquist, *Nature* **2016**, *533*, 73–76.
- [24] B. J. Reizman, K. F. Jensen, *Acc. Chem. Res.* **2016**, *49*, 1786–1796.

- [25] C. W. Coley, D. A. Thomas, J. A. M. Lummiss, J. N. Jaworski, C. P. Breen, V. Schultz, T. Hart, J. S. Fishman, L. Rogers, H. Gao et al., *Science* **2019**, 365.
- [26] D. C. Fabry, E. Sugiono, M. Rueping, *React. Chem. Eng.* **2016**, 1, 129–133.
- [27] V. Sans, L. Cronin, *Chem. Soc. Rev.* **2016**, 45, 2032–2043.
- [28] N. Vasudevan, E. Wimmer, E. Barré, D. Cortés-Borda, M. Rodríguez-Zubiri, F.-X. Felpin, *Adv. Synth. Catal.* **2020**.
- [29] R. Lee, *Chem. Ing. Tech.* **2019**, 91, 191–200.
- [30] R. Leardi, *Anal. Chim. Acta* **2009**, 652, 161–172.
- [31] a) D. C. Montgomery, *Design and analysis of experiments*, 8. Aufl., Wiley, Hoboken, NJ, **2013**; b) T. Lundstedt, E. Seifert, L. Abramo, B. Thelin, Å. Nyström, J. Pettersen, R. Bergman, *Chemom. Intell. Lab. Syst.* **1998**, 42, 3–40.
- [32] M. Baudin, M. Christopoulou, Y. Collette, J.-M. Martinez, “pyDOE”, can be found under <https://pythonhosted.org/pyDOE/>, **2017**.
- [33] U. Groemping, F. John, “R Commander Plugin for (industrial) Design of Experiments”, can be found under <https://cran.r-project.org/package=RcmdrPlugin.DoE>, **2014**.
- [34] a) B. Gleede, M. Selt, C. Gütz, A. Stenglein, S. R. Waldvogel, *Org. Process Res. Dev.* **2020**, 24, 1916–1926; b) M. Selt, R. Franke, S. R. Waldvogel, *Org. Process Res. Dev.* **2020**, 24, 2347–2355.
- [35] J. L. Röckl, M. Dörr, S. R. Waldvogel, *ChemElectroChem* **2020**, 7, 3686–3694.
- [36] S. Arndt, D. Weis, K. Donsbach, S. R. Waldvogel, *Angew. Chem. Int. Ed.* **2020**, 59, 8036–8041; *Angew. Chem.* **2020**, 132, 8112–8118.
- [37] J. du Boisberranger, J. Van den Bossche, L. Estève, T. J. Fan, A. Gramfort, O. Grisel, Y. Halchenko, N. Hug, A. Jalali, G. Lemaitre et al., “scikit-learn. Machine Learning in Python”, can be found under <https://scikit-learn.org/>.
- [38] “statsmodels. statistical models, hypothesis tests, and data exploration”, can be found under <https://www.statsmodels.org/>.
- [39] A. R. Katritzky, D. C. Fara, M. Kuanar, E. Hur, M. Karelson, *J. Phys. Chem. A* **2005**, 109, 10323–10341.
- [40] F. Strieth-Kalthoff, F. Sandfort, M. H. S. Segler, F. Glorius, *Chem. Soc. Rev.* **2020**, 49, 6154–6168.
- [41] a) G. dos Passos Gomes, R. Pollice, A. Aspuru-Guzik, *Trends Chem.* **2021**, 3, 96–110; b) K. Jorner, A. Tomberg, C. Bauer, C. Sköld, P.-O. Norrby, *Nat. Chem. Rev.* **2021**; c) A. J. Sterling, S. Zavitsanou, J. Ford, F. Duarte, *WIREs Comput. Mol. Sci.* **2021**.
- [42] A. F. Zahrt, S. V. Athavale, S. E. Denmark, *Chem. Rev.* **2020**, 120, 1620–1689.
- [43] Z. Obermeyer, E. J. Emanuel, *New Engl. J. Med.* **2016**, 375, 1216–1219.
- [44] K. T. Butler, D. W. Davies, H. Cartwright, O. Isayev, A. Walsh, *Nature* **2018**, 559, 547–555.
- [45] P. Pernot, F. Cailliez, *AIChE J.* **2017**, 63, 4642–4665.
- [46] a) G. Schneider, *Nat. Rev. Drug Discovery* **2010**, 9, 273–276; b) K. Artyushkova, S. Pylypenko, T. S. Olson, J. E. Fulghum, P. Atanassov, *Langmuir* **2008**, 24, 9082–9088; c) J. M. Serra, A. Corma, S. Valero, E. Argente, V. Botti, *QSAR Comb. Sci.* **2007**, 26, 11–26.
- [47] “scikit-optimize. Sequential model-based optimization in Python”, can be found under <https://scikit-optimize.github.io/stable/>.
- [48] “Tensorflow”, can be found under <https://www.tensorflow.org/>.
- [49] F. Häse, L. Roch, C. Kreisbeck, A. Aspuru-Guzik, “Phoenics”, can be found under <https://github.com/aspuru-guzik-group/phoenics>.
- [50] F. Häse, L. Roch, A. Aspuru-Guzik, “Gryffin”, can be found under <https://github.com/aspuru-guzik-group/gryffin>.
- [51] S. C. Larson, *J. Educ. Psychol.* **1931**, 22, 45–55.
- [52] a) M. Foscatto, V. R. Jensen, *ACS Catal.* **2020**, 10, 2354–2377; b) O. Lakuntza, M. Besora, F. Maseras, *Inorg. Chem.* **2018**, 57, 14660–14670.
- [53] D. Nematollahi, A. A. Taherpour, S. Jameh-Bozorghi, A. Mansouri, B. Dadpou, *Int. J. Electrochem. Sci.* **2010**, 5, 867–879.
- [54] R. Möckel, E. Babaoglu, G. Hilt, *Chem. Eur. J.* **2018**, 24, 15781–15785.
- [55] Y. Chen, B. Tian, Z. Cheng, X. Li, M. Huang, Y. Sun, S. Liu, X. Cheng, S. Li, M. Ding, *Angew. Chem. Int. Ed.* **2020**, 60, 4199–4207; *Angew. Chem.* **2021**, 133, 4245–4253.
- [56] C. Gütz, A. Stenglein, S. R. Waldvogel, *Org. Process Res. Dev.* **2017**, 21, 771–778.
- [57] D. E. Blanco, B. Lee, M. A. Modestino, *Proc. Natl. Acad. Sci. USA* **2019**, 17683–17689.
- [58] R. Hardian, Z. Liang, X. Zhang, G. Szekely, *Green Chem.* **2020**, 22, 7521–7528.
- [59] a) C. W. Coley, W. H. Green, K. F. Jensen, *Acc. Chem. Res.* **2018**, 51, 1281–1289; b) J. Liu, J. Li, W. Li, J. Wu, *ISPRS, J. Photogramm. Remote Sens.* **2016**, 115, 134–142.
- [60] B. J. Shields, J. Stevens, J. Li, M. Parasram, F. Damani, J. I. M. Alvarado, J. M. Janey, R. P. Adams, A. G. Doyle, *Nature* **2021**, 590, 89–96.

Manuscript received: March 11, 2021

Revised manuscript received: April 18, 2021

Accepted manuscript online: April 19, 2021